

# Exploring the Limits of the Data-Model-Theory Synergy: “Hot” MW Transitions for Rovibrational IR Studies

Xinchuan Huang (黄新川),<sup>1</sup> David W. Schwenke,<sup>2</sup> and Timothy J. Lee<sup>3</sup>

*Submit to J. Mol. Structure, the special issue Honoring Dr. Jon Hougen.*

---

<sup>1</sup> Corresponding Author: Xinchuan.Huang-1@nasa.gov

<sup>2</sup> Email: David.W.Schwenke@nasa.gov

<sup>3</sup> Corresponding Author: Timothy.J.Lee@nasa.gov

## Abstract

In order to further improve the accuracy of rovibrational IR line lists generated from the “Best Theory + Reliable High-resolution Experiment” (BTRHE) strategy from 0.01-0.05  $\text{cm}^{-1}$ , or 300-1500 MHz, to  $\sim$ 10 MHz, we explore the current limits of the Data-Model-Theory synergy by examining the accuracy and consistency of existing data, then propose that “hot” bands in microwave (MW) spectra is the solution we need for future enhancements. The Ames  $\text{SO}_2$   $J=0$ -20 rovibrational energy levels computed on the semi-empirically refined Ames-2 potential energy surface (PES) are fit to the Effective Hamiltonian (EH) model regularly used in the experimental infrared (IR) analysis for  $\text{SO}_2$  isotopologues. In the fitted EH(Ames) model, the rotational constants  $A/B/C$  and all 5 quartic centrifugal distortion constants display clear, systematic, and consistent patterns along the vibrational state energy or quanta. Such consistent patterns may facilitate the vibrational assignments for MW hot bands and extract more information from high temperature MW spectra. Some EH(Expt) analyses were carried out with the lowest order Coriolis Coupling term,  $C^1$ . Their constants should not be directly compared with other EH(Expt) and EH(Ames) results. After excluding them, our  $\delta = \text{EH(Ames)} - \text{EH(Expt)}$  analyses for 5 isotopologues (626, 636, 646, 628 and 828) indicates some loss of accuracy and consistency starting from vibrational states as low as  $2\nu_2$  or  $1000 \text{ cm}^{-1}$ . Some EH parameters, e.g.  $\delta_k$ , may have relative deviations as large as 50-100% and totally lose any recognizable patterns. This simply means that current EH(Expt) models do not have the system-wide consistency we need to further refine the EH(Ames) and Ames rovibrational IR line lists. A large part of such defects are probably inherited from the limited precision of experimental line positions, i.e.  $1\text{E-}3 \sim 1\text{E-}4 \text{ cm}^{-1}$ , or 3-30 MHz. This is confirmed in a series of truncation tests using the Ames data. Although the EH(Ames) consistency may help identify unreliable rovibrational EH(Expt) parameters, and make reliable predictions for minor isotopologues and unobserved vibrational bands, we believe only the highly accurate “hot” MW transitions can provide *real* enhancements for EH(Expt) accuracy and consistency. “Hot” MW spectra should play a more significant role in the future synergy of Data, Model, and Theory in the field of rovibrational IR studies.

## 1. Introduction

Modern atmospheric studies and infrared (IR) / radio astronomy need high quality opacity line lists for molecules playing important roles from remote sensing to life sustaining environmental characterization, etc. [1,2] It has been a trend to combine the advantages of high-resolution experimental analysis with computational spectroscopy to generate global opacity line lists suitable for high temperature, high energy studies. This synergy cycles amongst theory, modeling, and experiments and is mutually beneficial as it uses the strengths of each to overcome the shortcomings of all three. It has been a common strategy using experimental data to refine the theoretical computations. We call this approach the “Best Theory + Reliable High-resolution Experimental Data”, or BTRHE, which has led to  $0.01\text{-}0.05\text{ cm}^{-1}$  accuracy for rovibrational spectroscopy and >90% accuracy for intensity predictions.[3,4,13,5–12]

The question being addressed in this study is “How to push the BTRHE accuracy beyond  $0.01\text{ cm}^{-1}$  or 300 MHz?”. Recently we reported that the microwave (MW) prediction accuracy of minor  $\text{SO}_2$  isotopologues contained in the BTRHE-based Ames  $\text{SO}_2$  IR line lists can be further improved by two orders of magnitude, e.g. from  $\sim 1$  MHz to  $0.01\text{--}0.02$  MHz (A/B/C) or from hundreds of MHz to 0-10 MHz in the range of  $J<10\text{--}20$ .[14][Cite our paper here!] We fit Ames IR line lists to the same type of Effective Hamiltonian (EH) model that experimentalists used to analyze  $\text{SO}_2$  spectra, then compared the fitted EH(Ames) parameters vs. published experimental parameters, EH(Expt). Their differences are semi-linearly or quadratically correlated with the isotope substitutions or mass changes. Such mass dependencies allowed us to refine the line lists computed for minor isotopologues. Essentially, it is a 2<sup>nd</sup>-order implementation of the BTRHE strategy. Its realization requires reliable, accurate EH(Expt) parameters and high-quality theoretical line lists exhibiting superior consistency throughout all isotopologues.

That study focused on the purely rotational transitions (MW) of the vibrational ground state (GS). We proposed that the idea and implementation can be extended to vibrational excited states and we may improve the prediction accuracy for rovibrational IR lines by using a similar approach. Can we really do that? If so, how much better can we do for high energy bands? Taking  $\text{SO}_2$  as an example, this study investigates the possibility to improve the rovibrational line positions. We start by examining the parameter accuracy and consistency of EH(Ames) compared to existing high-resolution EH(Expt) and, mainly for those un-perturbed bands, we try to find a way to reach  $0.001\text{--}0.0001\text{ cm}^{-1}$  (or 3-30 MHz) prediction accuracy for rovibrational IR transitions, at least partially. Effectively, this process explores the limits of potential synergy between experiment, modeling and theory. This investigation will also help to answer more questions about the purely rotational transitions occurring within excited vibrational states. Can we extract useful information about the corresponding highly excited vibrational states from the “hot” bands observed in the MW region? Can we improve the “hot” band prediction accuracy to  $0.01\text{--}0.02$  MHz for A/B/C and 0-10 MHz for line positions?

The next section briefly introduces the data and methods adopted in this work. The results are presented and discussed in the following section. A summary is given in the Conclusions section.

Four Notes:

(1) Although many of the basic points and ideas reported in this work have been known for a while, we believe this is the first systematic or quantitative study of such carried out at  $1\text{E-}7\text{ cm}^{-1}$  or  $0.003\text{ MHz}$  precision level, especially for the BTRHE line lists;

(2) We sincerely appreciate all the experimentalists’ and modelers’ work and data quoted / used / compared to in this work. Without reliable high-resolution data and established EH models, the BTRHE IR line lists and

the 2<sup>nd</sup>-order BTRHE improvement would not exist. Any data issues mentioned in the EH(Expt) data are only general comments, not specific to any study, scientist, group, or molecule.

(3) The SO<sub>2</sub> isotopologues are denoted with the unit mass number of their O-S-O atoms. The main isotopologue <sup>32</sup>S<sup>16</sup>O<sub>2</sub> is denoted 626, and a mixed isotopologue <sup>17</sup>O<sup>34</sup>S<sup>18</sup>O is 748. Experimental data for 5 isotopologues are compared: <sup>32</sup>S<sup>16</sup>O<sub>2</sub> (626), <sup>34</sup>S<sup>16</sup>O<sub>2</sub> (646), <sup>33</sup>S<sup>16</sup>O<sub>2</sub> (636), <sup>16</sup>O<sup>32</sup>S<sup>18</sup>O (628), and <sup>32</sup>S<sup>18</sup>O<sub>2</sub> (828). Further, the vibrational states are labeled by their vibrational quanta  $\underline{v_1v_2v_3}$  for the symmetric stretch, bending, and anti-symmetric stretch, respectively. For example, 160 for  $v_1+6v_2$ , and 231 for  $2v_1+3v_2+v_3$ . The vibrational labels are underscored to avoid confusion with isotopologue labels. For convenience, the vibrational states of asymmetric isotopologues, i.e. 627 and 628, are labeled in a similar way. Except for the GS state, no EH(Expt) data was available for <sup>16</sup>O<sup>32</sup>S<sup>17</sup>O (627).

(4) Regular  $\Delta$  or  $\delta$  stands for differences and deviations, while italic  $\Delta$  and  $\delta$  are reserved for the quartic centrifugal distortion constants. For example,  $\delta(\delta)$  represents the change or difference of the distortion constant  $\delta$ .

## 2. Data, Method and Miscellaneous Concerns

### 2.1 EH(Expt) and EH(Ames) data

The experimental EH model parameters, i.e. EH(Expt), were taken from dozens of research articles published from 1984 to 2018, including those behind HITRAN[15–17], CDMS[18,19] and recent upgrades: refs.[20,21,30–35,22–29] for 626, refs [25,36–38] for 636, refs.[24,39–43] for 646, refs.[25] for 627, refs.[24,44–49] for 628, and refs.[42,44,50–52] for 828. We thank all contributors to this work including Flaud, Lafferty, Ulenikov and their colleagues. Inevitably, some states have more than one EH(Expt) available. If they are very similar in the comparison scale, only one EH(Expt) is used. Otherwise both (or all) are kept, if we have no way to determine whether *any* parameter is 100% “accurate” or “wrong”.

The EH(Expt) models were based on observed transitions. The transition datasets vary by band, measurement precision, spectrometer calibration, or the least-squares fitting weights chosen by each research group, etc. They all affect the accuracy and the consistency of EH(Expt) models. To maintain the highest consistency possible, the SPFIT[53,54] analyses use rovibrational energy levels for EH(Ames), instead of transitions. The  $J=0$ -20 Ames levels were computed on the Ames-2 PES[11], which is a high-quality *ab initio* potential energy surface (PES) semi-empirically refined with selected HITRAN data.[9,15,16] The accuracies of Ames-2 based line lists can be found in refs.[9,11–14]. Although the rovibrational levels come with 0.01-0.02 cm<sup>-1</sup> (300-600 MHz) deviations, they do have high consistency across isotopologues, as we recently reported[14]. Such high consistency should also exist for the EH(Ames) parameters fitted for vibrational states. All the Ames rotational levels were computed with 1E-8 cm<sup>-1</sup> or 0.3 kHz precision.

The EH(Ames) analyses were carried out on the lowest 80 states of 626, or the lowest 40 states of 627, 628, 636, 646 and 828. The EH(Ames) parameters for every vibrational state are fitted from its 441 rotational levels,  $J/K_a = 0$ -20, uniformly weighted. This means we always include those  $J/K_a/K_c$  levels with zero spin statistical weights so that the EH(Ames) parameter sets are consistent for all SO<sub>2</sub> isotopologues. The fitting was carried out using the SPFIT program[53,54] with the *A*-reduced Watson Hamiltonian conventionally adopted in SO<sub>2</sub> EH(Expt) analyses, e.g. see Refs.[18,32,37,55]. It includes band origins  $G_v$  ( $J=0$ ), rotational constants  $A/B/C$ , quartic centrifugal distortion constants ( $\Delta_k, \Delta_J, \Delta_{JK}, \delta, \delta_k$ ), and 17 sextic and higher order terms. See supplementary files SPFIT.files.xxx.tgz. The data uncertainty specified in SPFIT input varies from  $\pm 1E-4$  cm<sup>-1</sup> to  $\pm 1E-7$  cm<sup>-1</sup>, i.e. from 3 MHz to 3 kHz. A line is excluded as an “outlier” from the fit if its line position error is

larger than 5 times the uncertainty. Please note that all the rotational and quartic order constants have high S/N ratios, while the sextic and higher order constants may require higher  $J/K_a$  and/or  $1E-8\sim1E-9\text{ cm}^{-1}$  fits to reach good S/N ratios.

Coriolis Couplings and rovibrational resonances become so significant that many states  $> 3000\text{ cm}^{-1}$  are strongly perturbed and cannot achieve a successful fit with  $\sigma_{\text{RMS}} \leq 1E-4\text{ cm}^{-1}$  or 3 MHz. With a reduced data set, e.g.  $J=0\text{-}10$  or 16, some less-perturbed states can still obtain reasonably good EH(Ames) parameters such as  $A/B/C$ . But some other states cannot. Due to the high sensitivity of resonance couplings, the related EH terms are specific to the involved states and the isotopologue. This makes state-to-state extrapolations or isotopologue-to-isotopologue predictions very difficult, if not completely impossible. Fortunately, today most experimentally studied states are not perturbed, or very minimally perturbed, at least in the range of  $J\leq20$ . In practice, if the percentage of outliers is less than  $\sim10\%$  at  $1E-7\text{ cm}^{-1}$  (3 kHz) level, the state is considered "unperturbed" so its EH(Ames) constants are converged and reliable. Effectively, for the EH(Expt) vs EH(Ames) comparisons, only unperturbed state data are used. Heavily perturbed states reveal their outliers in the plotting analysis performed for the EH(Ames) parameters.

A consistency analysis on the unperturbed states will define the upper limits of any possible refinement. The extra difficulty in modeling perturbed states would introduce noise into the systematic pattern of fitted EH parameters, or eventually make it impossible to get any reasonable extrapolation or interpolation in the phase space of isotope substitution or vibrational quanta.

## 2.2 Method

The idea of refining EH(Ames) parameters with available highly accurate EH(Expt) models requires superior consistency for both. In Section 3, consistency checks within the  $X_{\text{Ames}}$  terms of EH(Ames) are performed to illustrate the vibrational progression of *lower* order parameters, where  $X = A/B/C$  or quartic centrifugal distortion constants. After establishing the EH(Ames) consistency, we examine the EH(Expt) consistency with  $\delta = X_{\text{Ames}} - X_{\text{Expt}}$ . The magnitude, patterns (if any), and uncertainties of  $\delta$  and  $\delta\%$  are analyzed.

We had tried to further investigate the consistency of EH parameter variations using a series of least-squares fits with 3 vibrational mode quanta. The next section presents general discussion about the results we acquired from such "fit-on-fit" analysis, but the current noise and uncertainties prevent us from obtaining definitive patterns. Although specific numerical details are not given, this is the major direction we will keep working on in the future. To be specific, here is the procedure used in "fit-on-fit". For each isotopologue, 8 separate fits are carried out for the 8 lower-order terms in EH(Ames) and EH(Expt). The expansion polynomials may have 1<sup>st</sup> to 4<sup>th</sup> order terms, depending on the available data and/or fitting results, e.g. one full 2<sup>nd</sup>-order fit on the 40  $B$ 's of 646, another full quartic fit on the 80  $\Delta$ 's of 626. Some strongly perturbed states appear as outliers which are rejected from the fits. Some outliers are common among the isotopologues, but some are specific to one isotopologue. This "fit-on-fit" analysis is similar with the Dunham expansion for vibrational state energies, but with  $X_0$  instead of  $X_e$ , and  $v_i$  instead of  $(v_i + \frac{1}{2})$ . For example, to determine EH parameter  $X_0$  and expansion coefficients  $C_x$  the following formula is used

$$X(v_1, v_2, v_3) = X_0(\text{GS}) + \sum_i C_{i,X} v_i + \sum_{ij} C_{ij,X} v_i v_j + \sum_{ijk} C_{ijk,X} v_i v_j v_k + \sum_{ijkl} C_{ijkl,X} v_i v_j v_k v_l + \dots \quad \text{Eq.(1)}$$

where  $X = A/B/C$  or  $\Delta_K/\Delta_{JK}/\Delta_J/\delta_J/\delta_K$ , and vibrational mode number  $i \leq j \leq k \leq l = 1, 2, \text{ or } 3$ . Next by comparing the sum of fitting residual squares, the fitted  $X_0$  and  $C_i, C_{ij}, \dots$  for different isotopologues, we should be able to gain more insight on the consistency of the current EH(Expt) and EH(Ames) data. Potentially it may enable a new round of fitting analysis: a fit on the fit-on-fit. In Ref.[14], we showed that the  $\delta(\Delta_K)$  is quadratically correlated

with the sum of O isotope mass inverses. For example, it would be interesting to see how the  $B_0$  and  $C_{i\_B}$  respond to isotope mass changes. Such an analysis can be done for the 8 lower order EH parameters, if necessary.

Ideally, if two sets of fitted EH models have similar consistency, their difference  $\delta$  should also be systematic. For example,  $X_{Ames}$  vs  $X_{Expt}$ ,  $X_{ISO1}$  vs  $X_{ISO2}$ , and  $X_{2v_2}$  vs  $X_{3v_2}$ , etc. In reality, many factors will have a non-negligible impact affecting the consistency. A major difference between experimental and Ames data is the precision. Most  $SO_2$  rovibrational line positions possess  $1E-4 \sim 1E-3 \text{ cm}^{-1}$  resolution, or 3-30 MHz, while the Ames data for low energy bands has at least  $1E-7 \text{ cm}^{-1}$  precision, or 3 kHz. How much will the EH(Ames) parameters change if the Ames energy levels are also truncated to  $1E-4 \text{ cm}^{-1}$  (3 MHz) or  $1E-5 \text{ cm}^{-1}$  (0.3 MHz) precision? SPFIT analyses are repeated with truncated Ames 626 data for those un-perturbed states. The fitting accuracy threshold is set equal to the truncation value ( $1E-4$  or  $1E-5 \text{ cm}^{-1}$ , 0.3-3 MHz), or to a tenth of the truncation value ( $1E-5$  or  $1E-6 \text{ cm}^{-1}$ , 0.03-0.3 MHz). Such “truncated” EH(Ames) are compared with the original EH(Ames) fitted at  $1E-7 \text{ cm}^{-1}$  level. Their differences are plotted and discussed in Section 3.

### 2.3 Planarity Linear dependency

Several issues need to be explained or clarified before we move into the technical details and results discussions. This is truly a learning process for us. One example is the linear dependency for the quartic distortion constants of planar molecules.

For planar molecules, there are only 4 linearly independent  $\tau_{abcd}$  constants.[56] The Watson A- or S- reduced Hamiltonian uses 5 quartic centrifugal distortion constants. This means the quartic terms are not completely unique or independent. The numerical solutions to such nonlinear least-squares problem will depend on the iterative solver and the relations among the constants. If they are correlated by symmetry rules, the dependency will not necessarily affect the stability and systematic pattern of the fitted constants. This is the case we found in an earlier study [14] using the EH(GS) model of 30  $SO_2$  isotopologues. The corresponding fitting uncertainty and constant uncertainty of  $\Delta/\delta$  are in the range of 0.0001 – 0.1%. Agreement with experimental data is also good, with systematic deviations that can be nicely approximated by linear or quadratic formula along a mass-inverse coordinate. These observations clearly indicate the EH(Ames) and EH(Expt) numerical analysis procedure may successfully reproduce the built-in symmetry relations among the  $\tau_{abcd}$  and  $\Delta/\delta$ . In other words, these constants are not really varying freely, because they are restrained by certain rules. Here is an example. If we use two identical  $\Delta_k$  in SPFIT, the fit will simply diverge. Any  $\Delta_{k1} + \Delta_{k2} = \Delta_k$  combination will give identical numerical results. However, if the non-linear least-squares solver can enforce the ratio  $\Delta_{k1}/\Delta_{k2} = 1$ , the fit will quickly converge to  $\Delta_{k1} = \Delta_{k2} = \Delta_k/2$ . In SPFIT, a negative IDPAR code will fix the ratio of this parameter to the previous parameter to a pre-determined value. Note the quality of fitted results is also dependent on the data precision. The sextic distortion constants also have special relations for planar molecules.[57]

### 2.4 spin-forbidden states

To maintain the best consistency between isotopologues and vibrational states, we use the complete sets of  $J/K_a=0-20$  energy levels computed on the refined  $SO_2$  PES. In our Ref.[14] study for the ground state, we investigated the EH(Ames) difference caused by excluding the nuclei spin forbidden levels which has statistical weight = 0. The  $\delta(\Delta_k)$  difference at  $1E-7 \text{ cm}^{-1}$  precision is less than 1-3 Hz. Hence, this will not impact our analysis. More importantly, using the full set can give relatively more reliable higher order terms.

### 2.5. EH analysis programs and their compatibility.

Many experimental IR/MW groups have their own programs for an EH analysis. These programs follow the same or similar theoretical background, but they were independently developed. Their implementations are not necessarily 100% equivalent. The treatment for resonance or coupling terms also may vary. Usually for the same state or polyad, most of the EH constants reported by different groups using different programs with different experimental datasets are very close to each other. But not all the results can be exactly reproduced. Please bear in mind that the SPFIT program we use for this study is not identical to the program used by the Ulenikov group, nor the program used by Flaud, Lafferty, Blake and their co-workers.

## 2.6. The lowest order terms in Fermi Resonance and Coriolis Coupling.

If an EH(Expt/Ames) analysis includes the lowest order term of a Fermi Resonance,  $F_0$ , as an independent parameter, the corresponding band origins  $G_v$  for the interacting states will shift away from the values determined from a similar EH(Expt/Ames) analysis without  $F_0$ . One example is the 626  $v_1/2v_2$  band origins. Using Ref.[32] data, our  $G_v$ 's match the value reported in Ref.[58] to 0.00002 cm<sup>-1</sup>. This is normal. But the EH(Expt) model in Ref.[32] treats  $F_0$  as a parameter fixed at some value estimated on a potential function.[27] As a result, the reported  $G_v$ 's are off by +0.0272 cm<sup>-1</sup> ( $v_1$ ) and -0.0272 cm<sup>-1</sup> ( $2v_2$ ), respectively. The  $G_v$  differences are 1000 times larger than what one would expect between two EH model/programs treating  $F_0$  implicitly, not explicitly.

Should the  $G_v$ 's in Ref.[32] be directly compared to the  $G_v$ 's in Ref.[58]? No. They are like apples vs. oranges. Results from a model using  $F_0$  will appear as an outlier in the data pool of EH(Expt) models without  $F_0$ , and vice versa. Since we are dealing with unperturbed states or less impacted states, it's natural to align with EH(Expt) models excluding  $F_0$ . We are not claiming the other EH(Expt) models are wrong. They are just plainly different.

$F_0$  cannot be determined from an EH fit. It requires a theoretical estimate from a high quality potential surface and vibrational analysis. This means it lacks the accuracy of other EH constants. Different  $F_0$  values lead to different band origins. In Ref.[32], the  $F_0 = 1.78$  cm<sup>-1</sup> estimated on a potential function [27] for the  $v_1/2v_2$  resonance led to a  $\pm 0.0272$  cm<sup>-1</sup> shift. We do not know their calculation details, but in our Ames PES based Quartic force field (QFF) + vibrational 2<sup>nd</sup>-order perturbation (VPT2) analysis, the  $v_1/2v_2$  split caused by this Fermi resonance is larger by nearly one order of magnitude,  $\pm 0.227$  cm<sup>-1</sup>.

The same discussion may extend to the lowest order term for Coriolis Coupling,  $C^1$ . In the case of SO<sub>2</sub> 626, introducing  $C^1$  makes the rotational constant  $C$  of  $v_1/v_3$  shift away from the Ref.[58] values by about  $\pm 4.125$  MHz. See Table 4 of Ulenikov *et al* [32]. In contrast, the shifts predicted by our VPT2 analysis are  $\pm 5.112$  MHz. We do not need to discuss the reliability of  $C^1$  or the shift prediction. Just state a fact: the  $C$  constants of  $v_1/v_3$  reported in Ulenikov *et al* [32] should not be compared to those in Flaud *et al* [58] or the EH(Ames) reported in this work, because the EH analyses were done differently.

Similar  $C^1$ -related EH(Expt) models have been reported for following 828 states: 100, 020 and 001 [50], 110 and 011[51], 200, 101 and 002[42], 111, 210 and 021 [52]. And several 628 states: 100 and 001 [45], 110 and 011 [47], 200, 101 and 002 [46]. For these states, EH(Ames) are compared with other EH(Expt) data if available.

To further clarify the rotational constant discrepancy between EH(Ames) and  $C^1$ -related EH(Expt) models, we have carried out a special case study on the SO<sub>2</sub> 626  $v_1/v_3/2v_2$  triad. Our study report is included as a supplementary file. Interested readers can find all the technical details there. It also summarizes what we

know, what we have learned, and how to think about the correlation between  $C^1$  and rotational constants. For example, the Fig.R1 c) in the report shows the  $A/B$  constants may shift  $> 0.1$  MHz in our simulations. This means the  $A/B$  rotational constants are also affected. This roughly matches the magnitude of  $\delta B$  shifts reported in Table 4 of Ref.[32], approximately  $-0.14$  MHz. We noticed that our SPFIT analysis could not reproduce their  $\delta A = -0.43$  MHz for  $v_1$ . It might be a minor difference rooted in the programs. In summary, those EH(Expt) results are program/parameter dependent, they should not be included in the figure and text comparison.

The EH(Ames)  $G$ 's can be found in our supplementary files, e.g. B121.par.b3y7-C-0.0MHz. Our VPT2 output file so2.spectro.out is also included.

### 3. Observation & Discussion

#### 3.1 Vibrational Consistency of EH(Ames) 626 parameters

The rotational constants of  $\text{SO}_2$  626 have very systematic patterns for most states below  $4500 - 5500$   $\text{cm}^{-1}$ , see Fig.1. In the scale of Fig.1, the vibrational dependence appears highly consistent. To facilitate graph reading, colored arrows are added to illustrate the trends of  $+nv_{1/2/3}$ . Starting from the GS at  $0 \text{ cm}^{-1}$ , in plot a),  $A$  rises nearly linearly with  $+nv_2$ ; in plot b),  $B$  decreases with the  $+nv_1$  and  $+nv_3$  stretches, but is less sensitive to  $+nv_2$ ; in plot c),  $C$  decreases consistently along the increasing quanta; and the  $A/B/C$  constants of the 1n1 states are right in the middle of 2n0 and 0n2 states, i.e.  $v_1+nv_2+v_3$  is in the middle of  $2v_1+nv_2$  and  $nv_2+2v_3$ . The isolated  $A/B/C$  variations upon every  $v_1/v_2/v_3$  quanta increment are averaged and summarized in Table 1. The semi-linear decrease of the  $C$  constants can be explained by the approximate  $v_1 \approx v_3 \approx 2v_2$  relation existing for both energies (wavenumber) and  $C$  reductions. These 1<sup>st</sup>-order approximations are not bad, while we believe 2<sup>nd</sup> or higher order approximations can give more accurate  $A/B/C$  predictions for those observed vibrational states or minor isotopologues, if 2 or 3 mode coupling terms and higher order terms are included.

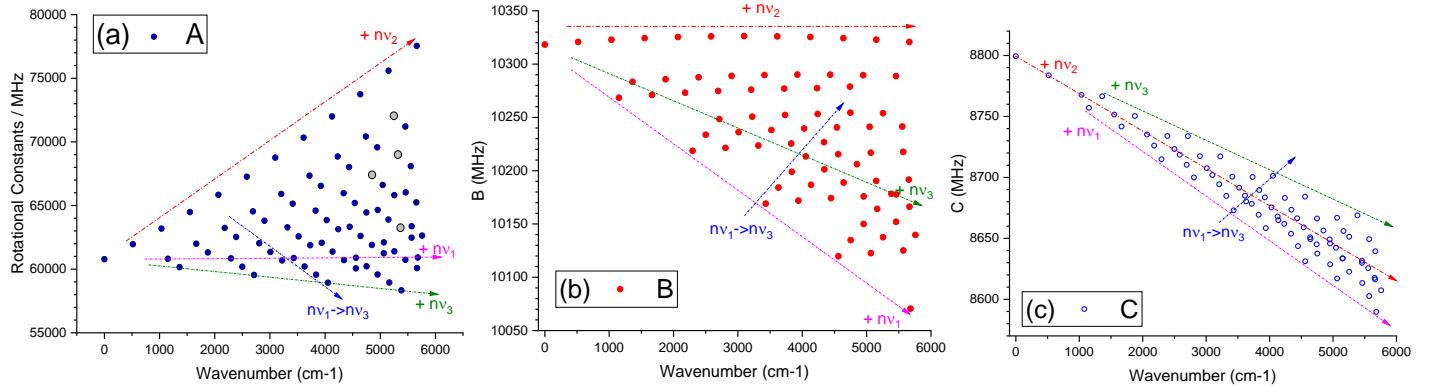


Fig.1 The vibrational dependence of  $A/B/C$  rotational constants of  $\text{SO}_2$  626, i.e.  $^{32}\text{S}^{16}\text{O}_2$ : a)  $A$ ; b)  $B$ ; c)  $C$ .

Table 1.  $A/B/C$  changes caused by one vibrational quanta increment in  $v_1/v_2/v_3$ , averaged from the  $A/B/C$  variations on unperturbed states shown in Fig.1.

(MHz)	$A$	$B$	$C$
$+v_1$	28	-49.5	-35
$+v_2$	1524	0 (-2 ~ 2)	-17
$+v_3$	-611	-35	-33

Only 4 out of the first 80 vibrational states are excluded from Fig.1. They are the 250 state at  $4846 \text{ cm}^{-1}$ , the 180 state at  $5252 \text{ cm}^{-1}$ , the 132 state at  $5354 \text{ cm}^{-1}$ , and the 260 state at  $5355 \text{ cm}^{-1}$ . The 132 and 260 states are in extremely strong resonance. But their un-perturbed  $A/B/C$  can be reliably predicted from the patterns

observed on other neighboring states. The gray circle examples in Fig.1a include a simple approximation  $A_{180} = A_{081} + A_{170} - A_{071}$ .

In Fig.2, the top row shows  $\Delta_K$ ,  $\Delta_{JK}$ ,  $\delta_K$  in MHz, while the bottom row is  $\Delta_J$  and  $\delta_J$  in kHz. A common feature in all 5 panels is that  $nv_1$  and  $nv_3$  have opposite effects on the quartic terms. In panel a), it is interesting that the  $\Delta_K$  data can be grouped by the sum of  $v_1/v_2/v_3$  quanta. See the two blue ovals for the sum = 4 or 7. In panel e), the  $nv_2$  and  $nv_3$  effects on  $\delta_J$  are similar and nearly parallel to each other. Along the  $nv_2$  series,  $\Delta_{JK}$  rises but the other 4 terms decrease.

Compared to A/B/C, the quartic centrifugal distortion constants are more sensitive to the numerical noise and resonance perturbations. The fitted EH(Ames) quartic terms start losing consistency on some states  $> 4000 \text{ cm}^{-1}$ . An exception is the  $\delta_J$  of 150 state at  $3719 \text{ cm}^{-1}$ . In panels c), d), and e), a few examples are given by black arrows pointing to the predicted location of their un-perturbed values. In d) and e), grey dots are put on the predicted locations for a few outliers that are out of range, not shown). For example,  $\Delta_J$  for 160 at  $4230 \text{ cm}^{-1}$ , and  $\delta_J$  for 330 at  $4955 \text{ cm}^{-1}$ .

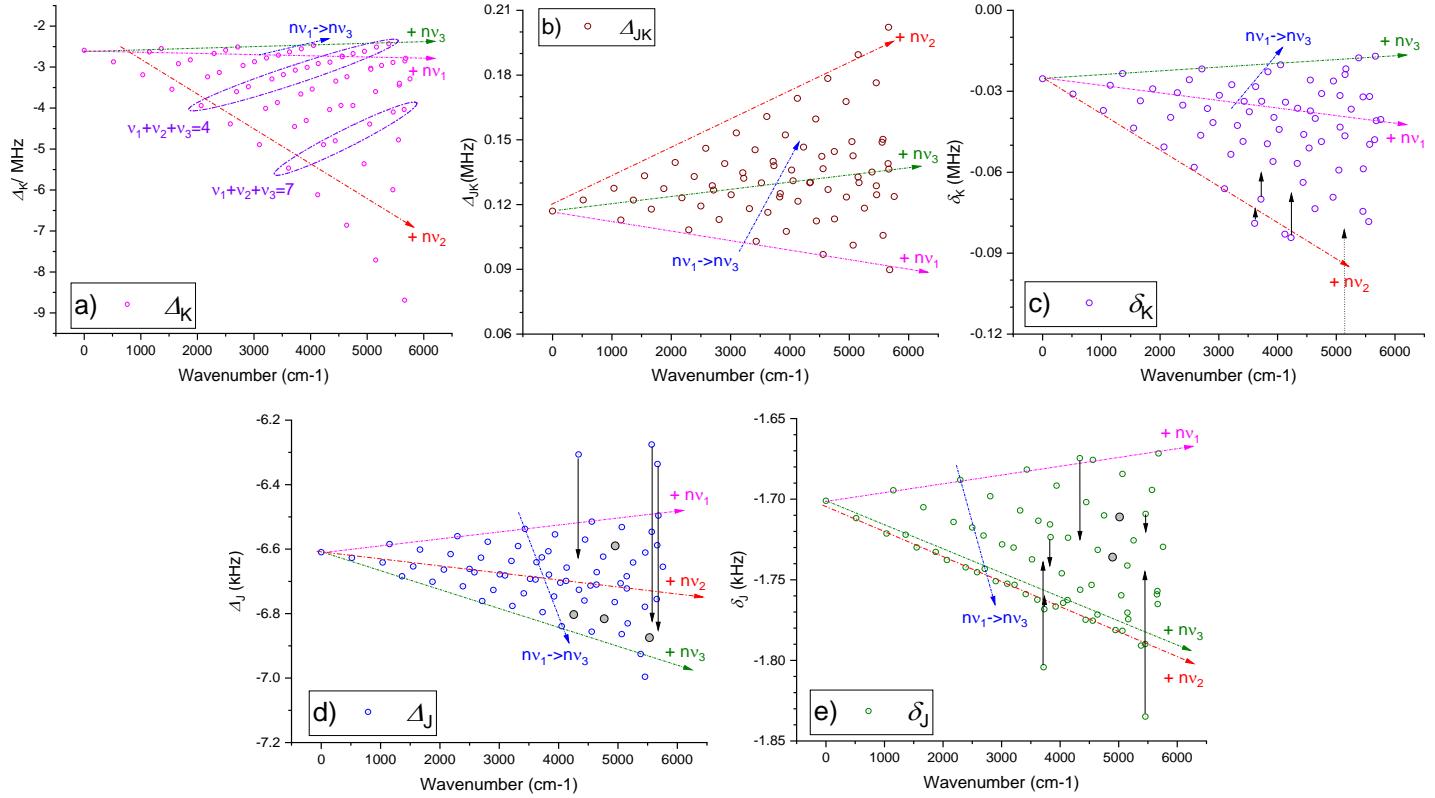


Fig.2 The vibrational dependence of Quartic centrifugal distortion constants of  $\text{SO}_2$  626, i.e.  $^{32}\text{S}^{16}\text{O}_2$ : a)  $\Delta_K$ ; b)  $\Delta_{JK}$ ; c)  $\delta_K$ ; d)  $\Delta_J$ ; e)  $\delta_J$

These observations confirm the BTRHE-based Ames  $\text{SO}_2$  IR lists do have enough precision for SPFIT analysis. There exist recognizable and consistent patterns for A/B/C and the quartic constants in the fitted EH(Ames). For higher energy states, more significant couplings and perturbations lead to larger relative deviations and more outliers in the quartic constants. On the other hand, the convergence defects in the original Ames rovibrational levels also rise along with energies, which will translate to larger errors and uncertainties in EH(Ames). Future refinements will require that all energies are tightly converged when reliable experimental data become available.

A test using seven 1<sup>st</sup> & 2<sup>nd</sup> order coefficients can reproduce most  $B_{Ames}^{626}$  within  $\pm 0.5$  MHz for 626 vibrational states up to 5700 cm<sup>-1</sup>. This means the perturbations are not very strong for most 626 levels. The EH(Ames) lower-order terms for the minor isotopologues 646, 636, 627, 628 and 828 have similar patterns and consistency. Their lowest 40 states are analyzed in this work (compared to 80 states for 626). For example, the simplest 1<sup>st</sup>-order approximation for  $C_{Ames}^{636}$  yields a pattern that uses 3 linear terms, which are different from the 626 values in Table 1:

$$C_{Ames}^{636}(v_1, v_2, v_3) \approx C_{Ames}^{636}(\text{GS}) - 41.2v_1 - 16.2v_2 - 33v_3, \text{ in MHz.}$$

This formula reproduces  $C_{Ames}^{636}$  values for all the 636 states up to 3500 cm<sup>-1</sup>, with an accuracy  $\pm 1.5$  MHz. Such qualitative tests can partially indicate the EH(Ames) consistency, provided that they are not overfitted.

Fortunately, most EH(Expt) models were reported for the unperturbed, or less perturbed states in the low energy region. Next we examine the  $\delta = X_{Ames} - X_{Expt}$  differences. Recall: those  $C^1$ -related EH(Expt) constants are excluded.

### 3.2 626 EH(Ames) – EH(Expt): absolute & relative differences

Fig.3a shows the  $\delta = X_{Ames} - X_{Expt}$  difference of rotational constants. The range of  $\delta A$  is  $-4 \sim +5$  MHz, no obvious patterns are found along energies or vibrational quanta. The range for most  $\delta B$  and  $\delta C$  is  $\pm 0.5$ -0.6 MHz. Only 1  $\delta C$  value for  $3v_2$ [34],  $C_{Ames} - C_{Expt} = 8751.52997 - 8752.4078 = -0.88$  MHz, is slightly larger but still acceptable. The average and  $\sigma_{\text{RMS}}$  of  $\delta_{A/B/C}$  for 21  $^{32}\text{S}^{16}\text{O}_2$  states are  $0.56 \pm 2.34$ ,  $0.30 \pm 0.16$ , and  $-0.26 \pm 0.18$  MHz respectively. If the  $3v_2$  related data is excluded, we have  $0.33 \pm 2.13$ ,  $0.29 \pm 0.16$  and  $-0.23 \pm 0.12$  MHz for the remaining 20 states.

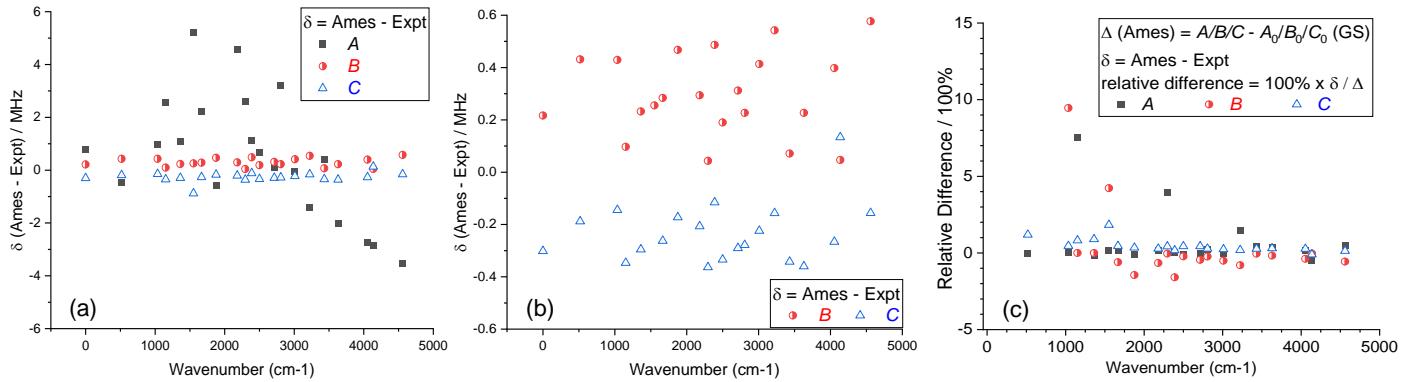


Fig.3 a) The  $\delta_{A/B/C}$  differences on 21  $\text{SO}_2$  626 vibrational states,  $\delta = X_{Ames} - X_{Expt}$ ; b) the  $\delta_B$  and  $\delta_C$  differences,  $\delta_C(3v_2)$  is out of range ; c) the relative differences as defined in text and the plot.

Correlation patterns of  $\delta B$  and  $\delta C$  are revealed in Fig.3b. It can be roughly formulated as  $\delta B - \delta B_{\text{GS}} \approx 2$  ( $\delta C - \delta C_{\text{GS}}$ ), and  $\delta B_{\text{GS}} \approx \delta C_{\text{GS}} + 0.5$ . One of exceptions is the state 211 ( $2v_1 + v_2 + v_3$ ) at 4137 cm<sup>-1</sup>: its  $\delta B$  is too low, and  $\delta C$  is too high. The  $\delta C(211) = 0.13$  MHz at 4137 cm<sup>-1</sup> is  $\sim 0.35$  MHz above the “predicted” range. This may be related to the fact that, for 211, only  $G_V$ ,  $A/B/C$  and  $\Delta_K/\Delta_{JK}$  were determined by fitting. The other parameters were fixed at predicted values during the fit.[29] Further investigation is beyond the scope of this work.

In general,  $\delta B$  and  $\delta C$  exhibit patterns along the stretches  $v_1$  and  $v_3$ , and the  $nv_1 \rightarrow nv_3$  series, but not so much for  $+nv_2$  series. The  $\delta$  differences in Fig.3 are less than 0.5% of the Fig.1 scale. Obviously, both EH(Expt) and EH(Ames) follow similar patterns and agree better than 99%. The recognizable pattern of  $\delta$  reflects their

differences, while the irregularities indicate inconsistency and uncertainty. For example, the  $\delta B - \delta B_{GS}$  and  $\delta C - \delta C_{GS}$  are close to zero MHz for the  $nv_1$  states, then rise along the  $nv_1 \rightarrow nv_3$  series. This suggests the rotational structure of the  $nv_1$  states in the Ames line list is more stable (or reliable) than that of  $nv_3$ . Since  $B$  and  $C$  follow an approximate linear correlation in Figs.1b and 1c, it is expected to see a linear correlation for their  $\delta$ 's and the difference between their  $\delta$ 's, e.g.  $\delta B - \delta B_{GS} \approx k(\delta C - \delta C_{GS})$ . But why  $k \approx 2$  in Fig.3b? This may be just accidental, or this may be related to the quality of the Ames PES.

It should be noted that ignoring a Coriolis coupling or a Fermi resonance may lead to errors in band origins and rotational constants, if the low  $J/K_a$  range is affected. The magnitude of the error is proportional to energy differences between the two states. As the comparison is mainly on less-impacted states, we believe only a small part of the differences, e.g.  $\sim 0.1$  MHz in Fig.3, results from ignoring such terms.

The uncertainty and inconsistency of  $\delta = X_{Ames} - X_{Expt}$  needs to be small enough so that we can use the patterns to calibrate EH(Ames) and improve BTRHE predictions, especially for those unobserved (or dark) bands. In Fig.3a, we estimate that the magnitude of noise for  $\delta A$  is about 2 MHz. The  $\delta B$  and  $\delta C$  noise for most states is probably within  $\pm 0.05$ - $0.15$  MHz. The main problem is defining and applying appropriate criteria to determine which  $\delta$  are outliers.

For now, one may suspect that the  $\delta A/\delta B/\delta C$  noise (or the inconsistency) comes from either EH(Expt) or EH(Ames), or both. The data currently available is not enough to draw any conclusion. It is also essentially impossible to separate the errors from the numerical noise in every piece of data. How do we know the BTRHE levels are *really* converged without *any* unknown/unexpected issues that would affect the EH(Ames) consistency? Experimentalists may carry out independent sets of measurement and analysis for a specific band using a different laboratory setup and programs, then to cross check the new EH(Expt) terms. For theory, we can increase the variational basis and check the convergence. The ultimate proof for the quality of EH(Ames) would be an totally independent BTRHE study generating a new PES, a new BTRHE line list, and a new EH(theory).

The “relative difference”  $\delta\%$  in Fig.3c is defined differently. Instead of using the original  $A/B/C$  constants as reference, it is computed with respect to  $\Delta X = X - X_0(GS)$ , where  $X = A/B/C$  and  $X_0$  is the constant value for the ground state GS. The  $X_0$  terms have been reliably determined for these isotopologues, e.g.[18,55]. To better determine the EH(Expt) and EH(Ames) difference patterns, it is more useful to compare with  $\Delta X$ . Consequently, the  $\delta\%$  distribution in Fig.3c is different from that in Fig.3a and 3b. For example,  $\delta B_{020} = 0.43$  MHz is totally normal in Fig.3b, but  $\Delta B_{020} = B_{020} - B_0(GS) = 4.5$  MHz, so the  $\delta\% \approx 10\%$ . This is one of the largest  $\delta\%$ . Most  $\delta\%$  are around or  $< 1\%$ .

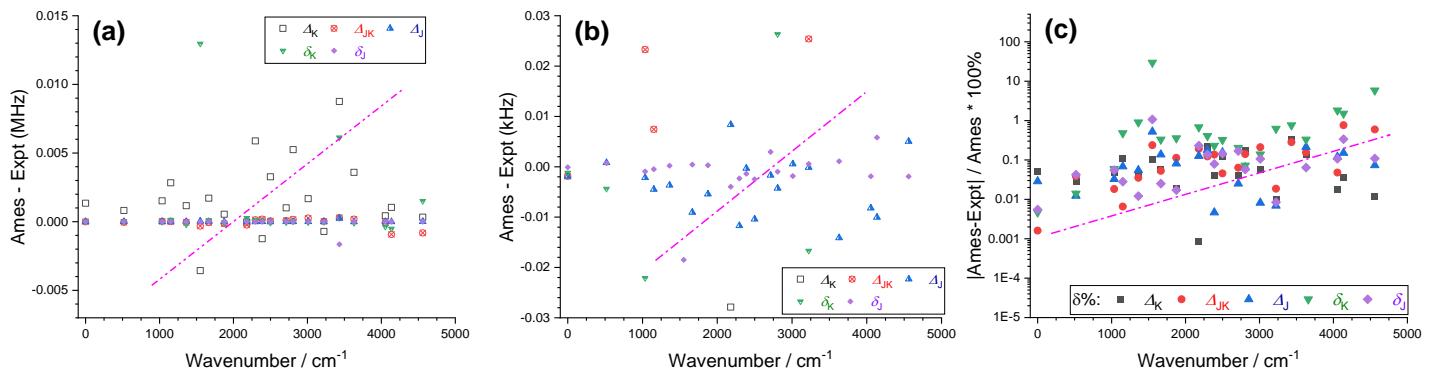


Fig.4 Ames vs Expt differences of five quartic centrifugal distortion constants of  $\text{SO}_2$  626 vibrational states: a)-b)  $\delta = X_{Ames} - X_{Expt}$ ; c) relative difference  $\delta\% = 100\% \times \delta/X_{Ames}$ .

Figs.4a-4b report the  $\delta$  differences of quartic centrifugal distortion constants in two different scales. Patterns along the  $nv_1-nv_3$  series are only partially recognizable for  $\Delta_K$ ,  $\Delta_{JK}$ ,  $\Delta_J$  and  $\delta_J$ , with a lot of noise and outliers. Fig.4b focuses on the  $\delta(\Delta_J)$  and  $\delta(\delta_J)$  constants, in the range of  $\pm 30$  Hz. The  $\delta(\delta_J)$  has only a barely recognizable pattern left. For  $\delta(\Delta_K)$  there is no recognizable pattern at all. For  $\Delta_K$  and states below  $4000$  cm $^{-1}$ , the  $\pm 0.005$  MHz disagreement translates into  $\sim 0.1\%$  relative differences. Compared to Fig.2, we are inclined to conclude that the current 626 EH(Expt) models do not have enough consistency or accuracy to guide the EH(Ames) refinement along vibrational quanta or wavenumbers. Again, the question is what would the *exact* patterns look like? Are their correlations linear or non-linear?

Figs.4c shows the absolute values of the relative  $\delta\%$  in log<sub>10</sub> scale, where the  $\delta$  values are divided by the corresponding EH(Ames) values. A general trend is that  $|\delta\%|$  for  $\Delta_{JK}$  (red),  $\Delta_J$  (blue),  $\delta_J$  (purple) and  $\delta_K$  (green) all rise along with energy. But the  $|\delta\%|$  for  $\Delta_K$  (black) seems relatively stable for the states  $< 4700$  cm $^{-1}$ . The majority of  $|\delta\%|$  is between 0.001% and 1%. Take  $\Delta_{JK}$  as example. The  $\delta\%(\Delta_{JK})$  rises from 1E-3% for the ground state to  $\sim 1\%$  at 4700 cm $^{-1}$ , so a three orders of magnitude increase. The  $\delta_K$  noise is the largest for the 5 quartic terms. This is expected. However, the largest exception  $|\delta\%|$  for  $\delta_K$  is found for the  $3v_2$  state at 1552 cm $^{-1}$ . We doublechecked the  $\delta_K$  stability in the EH(Ames) fits, then confirmed  $\delta_K = 0.0436$  MHz. The EH(Expt) model [34] reported  $\delta_K = 1.023E-6$  cm $^{-1}$ , or 0.0307 MHz. Such a large difference deserves careful investigation in a future study.

### 3.3 $\delta$ for 5 Isotopologues and Potential EH(Ames) Refinements

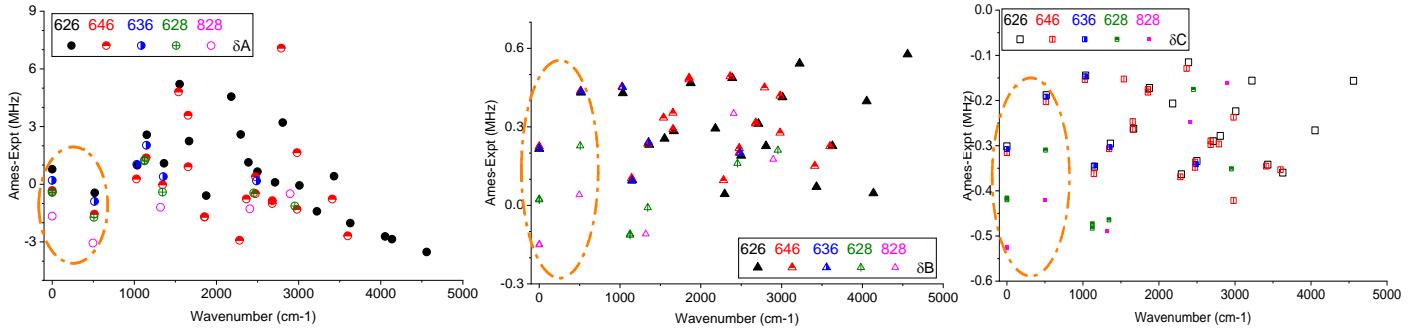


Fig.5 The  $\delta = X_{\text{Ames}} - X_{\text{Expt}}$  differences of A/B/C constants for 5 isotopologues: a) A; b) B; c) C.

To have a more inclusive estimate on the noise/consistency level for current EH(Expt) parameters, Fig.5 expands the  $\delta(A/B/C)$  comparison from 626 to 5 SO<sub>2</sub> isotopologues. In the 626  $v_1/v_3/2v_2$  case study we report in a supplementary file, the A/B constant deviations caused by the  $C^1$  term are found to be within  $\pm 0.20$  MHz. Deviations of such magnitude would show as noise in the panel b) and c) of Fig.5, but not significant in the scale of panel a). But for consistency, we exclude all the  $C^1$ -related EH(Expt) constants. The list of corresponding vibrational states is given in sec.2.6.

In addition, we identified two more EH(Expt) datasets that should not be included in our comparison. One is the 828 020 state [50] with  $\delta A = -266$  MHz. It was a dark state supporting the  $v_1/v_3$  analysis so its A/B/C were only predicted, not really fitted. The other is the 628 021 state [48], with  $\delta B = -64$  MHz,  $\delta C = 66$  MHz. It seems strange and we cannot explain. The differences between EH(Ames) and the theoretical prediction in Ref.[48] are just 0.3 MHz and 3 MHz, and no resonance or coupling terms were mentioned in Ref.[48]. As a result, there are 21/21/6/8/6 comparable EH(Expt) A/B/C datasets for 626/646/636/628/828, respectively. The mean  $\pm \sigma_{\text{RMS}}$  statistics for these 62 datasets are  $0.12 \pm 2.10$  MHz ( $\delta A$ ),  $0.24 \pm 0.18$  MHz ( $\delta B$ ), and  $-0.30 \pm 0.14$  MHz ( $\delta C$ ). It is consistent with the statistics for the 626 A/B/C constants reported in a previous section. This is

evidence for the isotopologue consistency in EH(Ames). One exception is the  $\delta A = 7.1$  MHz for the 646 210 state [59]. The  $\delta A$  for the 626 210 state [31] was only 3.3 MHz. Both EH(Expt) models had  $F_0$  but no  $C^1$ . This discrepancy may be investigated later.

In Fig.5, we focus on the ability to identify data outliers for both EH(Expt) and EH(Ames) and then how to use EH(Expt) to guide the EH(Ames) refinement. The  $\delta A$  and  $\delta B$  of the GS and  $v_2$  states of 5 isotopologues are marked by dashed circles. We can see consistent mass-dependency upon  $^{18}\text{O}$  and  $^{33/34}\text{S}$  substitution. The  $\delta B$  and  $\delta C$  patterns of 626-636-646 also match reasonably well in the whole range of plots b) and c). Such cross-isotopologue consistency further confirms the reliability of many EH(Expt) and EH(Ames) data sets, and suggests that the current  $\delta = \text{EH(Expt)} - \text{EH(Ames)}$  may help identify EH(Expt) outliers. On the other hand, the isotopologue consistency of EH(Expt) is suitable for some of the simplest EH(Ames) adjustments, but it is not good enough for a systematic refinement including high energy vibrational states, even if the refinements are limited to  $A/B/C$  only and the obvious outliers can be removed. For example, we can safely apply a 0<sup>th</sup>-order correction by adjusting all EH(Ames) lower-order terms  $X$  with the  $\delta$  we determined for a specific isotopologue in its ground state. In other words, the shifts (or the “refinements”) are independent of vibrational quanta. That is,

$$X_{\text{Predict}} \approx X_{\text{Ames}} - \delta X^{\text{GS}} = X_{\text{Ames}} - [X_{\text{Ames}}^{\text{GS}} - X_{\text{Expt}}^{\text{GS}}] = X_{\text{Expt}}^{\text{GS}} + X_{\text{Ames}} - X_{\text{Ames}}^{\text{GS}}$$

where  $X = A/B/C/\Delta_k/\dots$  etc. The reliability of this 0<sup>th</sup>-order fix will inevitably vary by isotopologue. It may reduce the EH(Ames) deviations for some states. However, it may not necessarily improve the overall agreement with EH(Expt). It depends on the sign and magnitude of shifts vs. the original  $\delta X$ .

Another simple adjustment is to apply the reliable  $\delta$  of a vibrational state ( $v_1, v_2, v_3$ ) determined for one isotopologue onto the same vibrational state of another isotopologue, if they have been predicted to be close enough. One such example is the 626-636-646 clusters in the  $\delta B$  and  $\delta C$  plots. As shown in Fig.5b and 5c, the differences between black (626), blue (636), and red (646) symbols for a given state are much smaller than the changes caused by the vibrational excitations. We may use a simple approximation such as:

$$X_{\text{Predict}}^{\text{ISO2}} = X_{\text{Ames}}^{\text{ISO2}} - [X_{\text{Ames}}^{\text{ISO1}} - X_{\text{Expt}}^{\text{ISO1}}] = X_{\text{Expt}}^{\text{ISO1}} + X_{\text{Ames}}^{\text{ISO2}} - X_{\text{Ames}}^{\text{ISO1}}$$

In this way, the prediction accuracy for many isotopologue bands can be significantly improved. Given an accurate  $X_{\text{Expt}}^{\text{ISO1}}$ , it is possible to reduce the  $X_{\text{Ames}}^{\text{ISO2}}$  errors by 90-95%, or even more. Scaling factors may be introduced to match the  $X_{\text{Expt}}^{\text{ISO2}}$ . But this approach only works if the isotope dependence of  $\delta$  is so small that it can be safely ignored, when compared to the original  $\delta$ . It does not work well for the 626-628-828 series. More importantly, we must first have an accurately determined  $X_{\text{Expt}}^{\text{ISO1}}$  for one isotopologue (ISO1), which itself is still a topic under investigation. Other questions people may ask include: How do we know if one  $X_{\text{Expt}}^{v_1 v_2 v_3}$  is really more reliable than another  $X_{\text{Expt}}^{v_1 v_2 v_3}$ ? No accidental clustering from bad data (either Expt or Ames)? etc. Sometimes it is possible to conclude one  $X_{\text{Expt}}^{v_1 v_2 v_3}$  set is more reliable than another, with crosscheck +  $\delta$  range + BTRHE consistency. Minor isotopologues with less abundance usually have relatively larger uncertainties in their EH(Expt) terms.

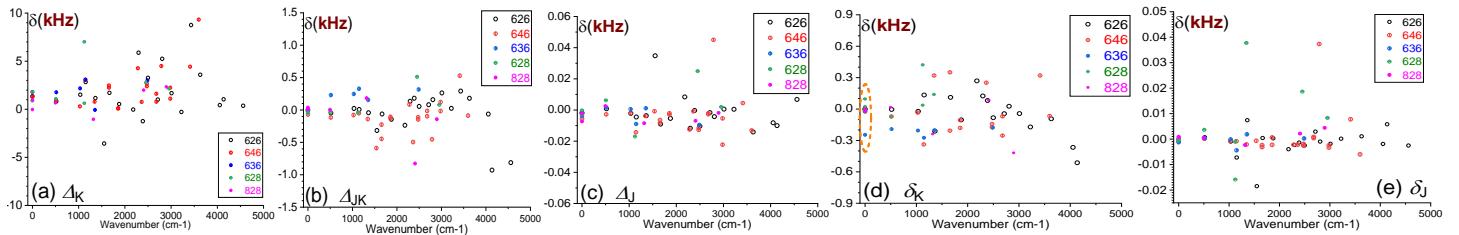


Fig.6  $\delta(X_{\text{Ames}}^{\text{ISO}} - X_{\text{Expt}}^{\text{ISO}})$  for the five quartic EH terms of 5  $\text{SO}_2$  isotopologues: a)  $\Delta_K$ ; b)  $\Delta_{JK}$ ; c)  $\Delta_J$ ; d)  $\delta_K$ ; e)  $\delta_J$  in kHz.

Fig.6 presents the isotopologue  $\delta$ s for the 5 quartic centrifugal distortion constant terms of the EH models for five isotopologues. As we show at the end of the study report in the supplementary file, the  $C^1$  term will cause  $\sim 1\%$  changes on  $\Delta_K/\Delta_J/\delta_J$  and even larger changes on  $\Delta_{JK}/\delta_K$ . All those EH(Expt) modes are excluded. It turns out that only the GS and the lowest  $v_2$  states have small uncertainties or noise in their  $X_{\text{Expt}}^{\text{ISO}}$ . As shown in Fig.2, the  $X_{\text{Ames}}^{\text{GS}}$  of minor isotopologues have recognizable, systematic patterns along vibrational quanta increments. The  $A/B/C$  and quartic terms of vibrationally excited states should also follow the approximate linear or quadratic correlation with the isotope substitutions, just as we reported for the GS in Ref.[14]. However, Fig.6 shows that there are not consistent changes along S or O isotope substitutions: black 626  $\rightarrow$  blue 636  $\rightarrow$  red 646, or black 626  $\rightarrow$  green 628  $\rightarrow$  magenta 828. For example, the blue 636 data in the Fig.6b plot for  $\delta\Delta_{JK}$  are  $\sim +0.25$  kHz larger than the black 626 data and red 646 data, which agree well with each other.

Similar to Fig.4, the  $\delta(\delta_K)$  in Fig.6d loses all recognizable patterns for vibrational states  $> 1000 \text{ cm}^{-1}$ . But the  $\delta(\delta_K)$  for the ground state is still a good example for EH(Expt) purity check. See the orange, dashed oval in Fig.6d. There are two  $\delta(\delta_K)$  for 628 (green), two  $\delta(\delta_K)$  for 828 (magenta), and two  $\delta(\delta_K)$  for 636 (blue). They should be very close to the  $\delta(\delta_K)$  of 626 (black) and 646 (red). Therefore, it is reasonable to assert that the experimental  $\delta_K$  data behind the 628  $\delta(\delta_K)=0.09$  kHz [24] and the 636  $\delta(\delta_K) = -0.25$  kHz [36] probably have more uncertainty or noise than the data behind 628  $\delta(\delta_K)=0.02$  kHz [44] and 636  $\delta(\delta_K) = 0.02$  kHz [18,55].

### 3.4. From “fit-on-fit” to MW+IR analysis

With EH(Ames), EH(Expt) and  $\delta$  available, we tried to fit their changes to the corresponding vibrational quantum numbers. Here is the Eq.1 mentioned in Sec.2.

$$X(v_1, v_2, v_3) = X_0(\text{GS}) + \sum_i C_{i,X} v_i + \sum_{ij} C_{ij,X} v_i v_j + \sum_{ijk} C_{ijk,X} v_i v_j v_k + \sum_{ijkl} C_{ijkl,X} v_i v_j v_k v_l + \dots \quad \text{Eq.1}$$

where  $X=A/B/C$  or  $\Delta_K/\Delta_{JK}/\Delta_J/\delta_K/\delta_J$ , and vibrational mode number  $i \leq j \leq k \leq l = 1, 2, \text{or } 3$ . Ideally, the fitted coefficients  $C_i, C_{ij}, C_{ijk} \dots$  can be used to predict the unperturbed  $X$  values for other bands. It is also possible to further fit the fitted  $C$  coefficients with respect to the isotope mass differences:

$$C_{\text{ISO}} - C_{626} = a_1 \Delta_0 + b_1 \Delta_S + a_2 \Delta_0^2 + b_2 \Delta_S^2 + \dots, \text{ where } \Delta_0 = \sum_{i=1}^2 \left( \frac{1}{m_{O_i}^{\text{ISO}}} - \frac{1}{m_{O_i}^{626}} \right), \Delta_S = \frac{1}{m_S^{\text{ISO}}} - \frac{1}{m_S^{626}}$$

This approach assumes smaller fitting deviations ( $\sigma_{\text{RMS}}$ ) indicate better consistency, which is not necessarily true in least-squares fit. Due to the unpredictable data quality, it is a tedious procedure to find the smallest  $\sigma_{\text{RMS}}$  with a given number of coefficients. The least-squares fitting procedure was carried out for available EH(Expt) data and the EH(Ames) data for the lowest 40-80 vibrational states. Using 4 to 21 coefficients, we tried linear, quadratic, cubic and up to quartic order fits. The  $\sigma_{\text{RMS}}$  will be less useful if the  $N_{\text{coe}}/N_{\text{data}}$  ratio is high, and the “best” coefficient set would rely on the structure of reliable “unperturbed” EH data. In addition, it is hard to appropriately separate the EH(Expt) noise from normal fitting errors and the resonance/coupling

perturbations. The outlier states vary from one isotopologue to another isotopologue, and from one parameter to another parameter. We have not been able to find a quantitative basis establishing the “best coefficient” set for the hundreds of sets of fitting results. Actually, we were lost in this “fit-on-fit” analysis. A new direction is needed, which should be a basic and straightforward one.

The only thing our “fit-on-fit” analysis has verified is what we already knew before running them: the EH(Ames) consistency are usually higher than EH(Expt) consistency by *one* order of magnitude. In other words, given the same set of vibrational states, the “fit-on-fit”  $\sigma_{\text{RMS}}$  for the EH(Ames) data is only 5-30% of the  $\sigma_{\text{RMS}}$  obtained from the fit on EH(Expt) terms. In a few cases they are similar and may be considered equal. From this aspect, it does provide some insights, but cannot really give us the answer we need. Our goal “to systematically refine EH(Ames) and IR list ***predictions*** with available EH(Expt) data” cannot be realized without knowing how exactly the *reliable and accurate* EH(Expt) terms would vary along the energy or vibrational quanta, *with high precision*. With the current uncertainties in the  $\delta$  of quartic order terms, the exact variation patterns cannot be determined, so we can only do one of the two simplest adjustments proposed in Sec.3.3. More advanced refinements inevitably need reliable vibrational quanta dependence and/or isotope substitution mass effects. None of these can be done without experimental data of higher precision.

Instead of trying to do the “mission impossible” estimation for the uncertainty inside dozens or hundreds of EH(Expt) models and thousands of parameters, it should be more realistic to improve the precision of experimental data and the consistency of fitted EH(Expt) models. Current MW and Sub-MM spectral techniques support line position accuracy as good as 1-10 kHz (0.3-3E-7 cm<sup>-1</sup>) or even better. But the IR broadening affected line width is on the level of 30-90 MHz (0.001 ~ 0.003 cm<sup>-1</sup>), which is 4 orders of magnitude larger. If one can measure the “hot” pure-rotational MW spectra of vibrationally excited states at either room temperature 296K or higher temperature, the EH(Expt) for that vibrational state would be a significant upgrade. If the number of measured transitions is less than what is needed for an independent EH model analysis, it can be fitted together with IR data with appropriate weights. Actually, this probably will be the best (and more realistic) way to go, because MW and IR data probably will have different  $J/K$  ranges, and the fitted MW based EH(Expt) model may fail to extrapolate reliably into the IR  $J/K$  range.

One perfect example is the  $^{32}\text{S}^{16}\text{O}_2$   $\nu_2$  study reported by Ulenikov *et al.* in 2018 [35]. They included 148 highly accurate MW transitions (taken from Ref.[26]) and 4232 IR transitions in their EH analysis, and reached  $\sigma_{\text{RMS}} = 20.1$  kHz for MW and 1.5E-4 cm<sup>-1</sup> (i.e. 4.5 MHz) for IR, respectively. These uncertainties have reached the experimental uncertainty level for both. If all the EH(Expt) models can be re-analyzed with similar accuracy, the Figs.3-6 may look different.

The MW band intensity and IR intensity need to be higher than the sensitivity threshold of spectrometers. Consider the purely rotational bands of 628  $\nu_2$  series as an example. At 296K, from  $\nu_2 \leftarrow \nu_2$  to  $5\nu_2 \leftarrow 5\nu_2$ , the maximum MW transition intensity drops from 2.2E-22 to 1.2E-26 cm/molecule. Every additional quanta reduces the intensity by one order of magnitude. The intensity of the peak transition  $13_{13,1} \leftarrow 12_{12,0}$  in the  $(h+1)\nu_2 \leftarrow n\nu_2$  series around 577 cm<sup>-1</sup> also decreases from ~3E-22 ( $2\nu_2 \leftarrow \nu_2$ ) to ~3E-26 ( $6\nu_2 \leftarrow 5\nu_2$ ) cm/molecule. Again, one order of magnitude less per quanta. For weaker or higher energy bands, higher temperature MW or Sub-MM spectra are needed.

Another benefit of “hot” MW+IR analysis is to prevent the multiple local minima phenomena when fitting thousands or tens of thousands of transitions.[14] Highly accurate line positions can provide critical constraints on EH parameters. See Ames SPFIT tests in the next section. Actually, if “hot” MW lines can maintain enough accuracy, a “hot” MW only study may also work. Although the discussion in Secs.3.2-3.3 are

mainly on the “Unperturbed” states, this approach may work for strongly coupled rovibrational bands, too. Primitive EH analysis can take the BTRHE line lists as input, figure out the necessary coupling terms, then use the fitted EH(BTRHE) as a starting reference to analyze the “hot” MW spectra. In the earlier stage of analysis, the approximate  $A/B/C$  constants can help identify the associated vibrational states (see Fig.1), then partition functions, then intensities, etc.

### 3.5 Precision Truncation Effects: Tests on EH(Ames)

From spectrometer calibration to the choice of line broadening profiles, and ground state combination difference (GSCD) analysis, etc., the experimental line positions are determined with certain uncertainties. Most  $\text{SO}_2$  rovibrational IR studies were reported with  $1\text{E-}4 \sim 1\text{E-}3 \text{ cm}^{-1}$  precision, or 3-30 MHz. Then the fitted EH(Expt) parameters are used to generate thousands of lines later collected in HITRAN[15,16] and other databases. We have concluded the current set of EH(Expt) models do not have enough consistency for systematic refinement on all EH(Ames) lower order terms, and the limited accuracy of IR transitions should be a main reason. We have proposed “Hot MW+IR” approach as a fix. Since the Ames rovibrational data we use in this study has (at least)  $1\text{E-}7 \text{ cm}^{-1}$  (3 kHz) precision, we may truncate its digits to simulate the precision loss and determine the corresponding magnitude of accuracy & consistency losses. Here we report the  $1\text{E-}4 \text{ cm}^{-1}$  (3 MHz) and  $1\text{E-}5 \text{ cm}^{-1}$  (0.3 MHz) truncation effects on the lower order EH(Ames) terms for unperturbed  $\text{SO}_2$  626 states. The results support our claim that “Hot” MW (+IR) can improve the EH(Expt) consistency with more accurate line positions.

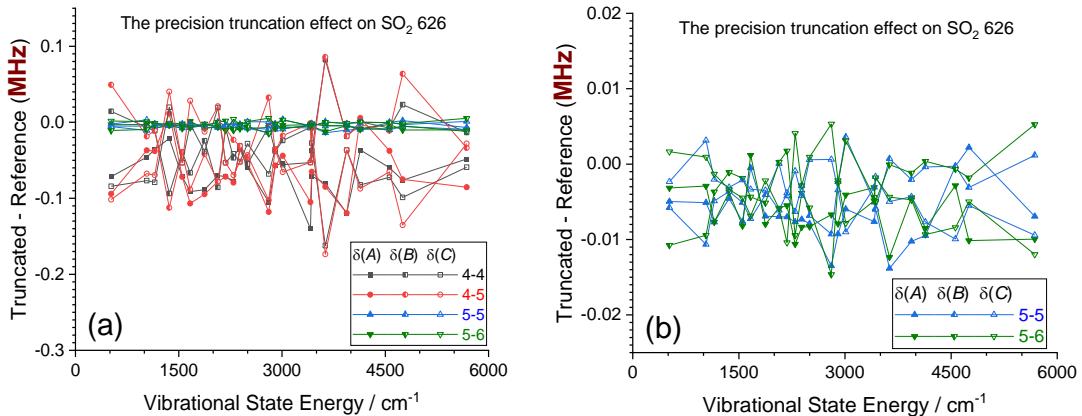


Fig.7 A/B/C differences between the EH(truncated) and EH( $1\text{E-}7$ ) reference values. a) four truncated sets in a range of 0.3 MHz; b)  $1\text{E-}5 \text{ cm}^{-1}$  (0.3 MHz) truncation differences in a range of 0.02 MHz.

From the ground state to  $5700 \text{ cm}^{-1}$ , 23 vibrational states are selected for the truncation test. This means more than 85-90% of their  $J=0-20$  rotational levels can be successfully fit with  $\delta \leq 5\text{E-}7 \text{ cm}^{-1}$  or 15 kHz. This is similar to the precision of MW data used in CDMS EH models. The corresponding EH(Ames) are taken as the “standard reference”. The SPFIT analysis is repeated 4 times for each MW line set: truncation to  $1\text{E-}4 \text{ cm}^{-1}$  + fitting deviations  $\delta \leq 5\text{E-}4 \text{ cm}^{-1}$  or  $5\text{E-}5 \text{ cm}^{-1}$ , truncation to  $1\text{E-}5 \text{ cm}^{-1}$  + fitting deviations  $\delta \leq 5\text{E-}5 \text{ cm}^{-1}$  or  $5\text{E-}6 \text{ cm}^{-1}$ . Difference between the resulting EH parameters and the “standard reference” values are shown in Fig.7 and Fig.8. They are denoted by the truncation level + the fitting accuracy level: 4-4 (black), 4-5 (red), and 5-5 (blue), 5-6 (green), respectively. The first number denotes the truncation level, while the second number indicates the fitting accuracy specified in SPFIT. For example, “4-5” means  $1\text{E-}4 \text{ cm}^{-1}$  truncation +  $\delta \leq 5\text{E-}5 \text{ cm}^{-1}$ , i.e. 3 MHz truncation +  $\delta \leq 1.5 \text{ MHz}$ .

In Fig.7a, the  $\delta(A/B/C)$  arising from the  $1\text{E-}4 \text{ cm}^{-1}$  truncation are in the range of  $-0.2 \text{ MHz} \sim +0.1 \text{ MHz}$ . In Fig.7b for the  $1\text{E-}5 \text{ cm}^{-1}$  truncation, the  $\delta(A/B/C)$  range is  $-0.015 \sim +0.007 \text{ MHz}$ , i.e. smaller by one order of magnitude.

Quartic terms also exhibit a similar order of magnitude loss in accuracy from 1E-5 to 1E-4  $\text{cm}^{-1}$ . See Fig.8a for the  $\delta$  values and Fig.8b for  $\delta\%$ . It is obvious the fitting accuracy level is a minor factor, while the truncation level is the primary one. If we push the fitting accuracy beyond the truncation level, the  $\delta$ s have larger oscillations. In other words, in order to achieve 0.01-0.02 MHz accuracy for the  $A/B/C$  constants, rovibrational IR spectral data need to have 1E-5  $\text{cm}^{-1}$  precision, or better.

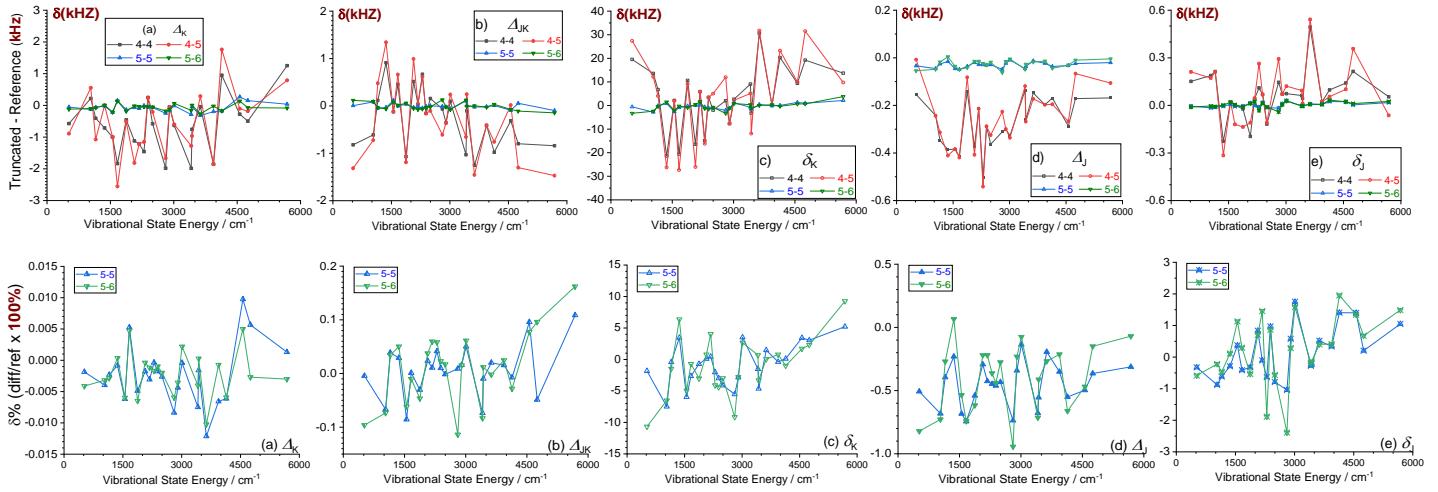


Fig.8  $\delta = X(\text{truncated}) - X(\text{reference 1E-7})$  differences (top row) and relative differences  $\delta\% = \delta / X(\text{reference 1E-7})$  (bottom row) for 5 quartic constants, from 4 truncation tests: a)  $\Delta_K$ ; b)  $\Delta_{JK}$ ; c)  $\delta_K$ ; d)  $\Delta_J$ ; e)  $\delta_J$

These analysis results are consistent with what we found in earlier sections. For example,  $\delta_K$  has the largest  $\delta\%$  and it is the most difficult quartic term to converge. At 1E-4  $\text{cm}^{-1}$  truncation, the magnitudes of  $\delta$  are  $\Delta_K \pm 2\text{kHz}$ ,  $\Delta_{JK} \pm 1\text{ kHz}$ ,  $\Delta_J \pm 0.3\text{ kHz}$ ,  $\delta_J \pm 0.4\text{ kHz}$ , and  $\delta_K \pm 30\text{ kHz}$ . The magnitudes of relative  $\delta\%$  are:  $\Delta_K \pm 0.05\%$ ,  $\Delta_{JK} \pm 1\%$ ,  $\Delta_J \pm 5\%$ ,  $\delta_J \pm 20\%$ , and  $\delta_K \pm 100\%$ . The bottom row of Fig.8 presents the 1E-5  $\text{cm}^{-1}$  truncation results only. The relative  $\delta\%$  have been reduced to:  $\Delta_K \pm 0.01\%$ ,  $\Delta_{JK} \pm 0.1\%$ ,  $\Delta_J \pm 1\%$ ,  $\delta_J \pm 2\%$ , and  $\delta_K \pm 10\%$ . Given a 1E-6  $\text{cm}^{-1}$  truncation, all relative  $\delta\%$  would be expected to be  $\pm 0.001\text{--}1\%$ , i.e. one more order of magnitude smaller.

To discuss the errors introduced by 1E-4  $\text{cm}^{-1}$  truncation, it is necessary to distinguish the errors on line positions and errors on energy levels. We randomly choose an example from the 828 ground state, the  $J_{\text{Ka},\text{Kc}} = 19_{9,10}$  level at 240  $\text{cm}^{-1}$ . A  $\pm 0.1\text{ MHz}$  change on  $A/B/C$  leads to 8/15/15 MHz difference on the level, a 2/1/0.3/0.4/30 kHz change on  $\Delta_K/\Delta_{JK}/\Delta_J/\delta_J/\delta_K$  leads to 13/31/43/2/2 MHz difference on the level, respectively. In Fig.6d, the average  $\delta(\delta_K) = 1.3\text{ kHz}$ , which would only lead to 100 - 200 kHz variation on energy levels, or 3-7E-6  $\text{cm}^{-1}$ . This explains the source of the  $\delta(\delta_K)$  noise: the IR line position precision is insufficient for a reasonable  $\delta_K$  fit. Two possible solutions are: a) either do the “Hot MW”(+IR) analysis, or b) adopt the EH(Ames)  $\delta_K$  values from Fig.2e with the GS difference correction.

Tables 2 and 3 report the statistics for the data in Fig.7 and Fig.8. We can see that all  $\delta(A/B/C)$  have their mean values  $< 0$ , which means the truncation has reduced the  $A/B/C$  values by 0.001-0.1 MHz. In Fig.7, the  $\delta_B$  and  $\delta_C$  data are approximately mirrored with respect to a horizontal line at -0.03 MHz (for the 1E-4  $\text{cm}^{-1}$  truncation) or -0.004 MHz (for the 1E-5  $\text{cm}^{-1}$  truncation). This is pre-determined by the Watson Hamiltonian formula used in the EH analysis.

Table 2.  $A/B/C$  constant differences in EH(Ames) fitted from precision truncation tests, compare to 1E-7  $\text{cm}^{-1}$  fit reference values. Mean  $\pm \sigma_{\text{rms}}$  are given in MHz unit. See text for denotations.

Mean $\pm \sigma_{\text{rms}}$ (MHz)	A	B	C
--------------------------------------	---	---	---

Average	61975	10239	8709
$\delta(4\text{-}4)$	-0.069±0.028	-0.022±0.037	-0.053±0.039
$\delta(4\text{-}5)$	-0.067±0.035	-0.019±0.046	-0.053±0.049
$\delta(5\text{-}5)$	-0.007±0.004	-0.004±0.003	-0.004±0.003
$\delta(5\text{-}6)$	-0.006±0.003	-0.003±0.005	-0.004±0.004

In the quartic term statistics (see Table 3), the truncation effect on  $\Delta_J$  is dominantly negative. The average truncation effects (mean value) are also negative for  $\Delta_K$  and  $\Delta_{JK}$ , while slightly positive for  $\delta_K$ . The reason needs further investigation. This could be molecule specific. A consequence of these changes is that the planarity-condition relation among the 5 fitted quartic constants will be less rigorous in the lower precision fits.

Table 3. Average values of Quartic centrifugal distortion constants of 23 “unperturbed”  $^{32}\text{S}^{16}\text{O}_2$  states, and the  $\text{mean} \pm \sigma_{\text{rms}}$  statistics for  $\delta X_{\text{Ames}} = X_{\text{Ames}}^{\text{truncated}} - X_{\text{Ames}}^{1\text{E}-7}$  differences, in MHz. See text for denotations.

	$\Delta_K$	$\Delta_{JK}$	$\delta_K$	$\Delta_J$	$\delta_J$
Average (MHz)	2.989	-0.1198	0.03523	0.00663	0.00171
Mean± $\sigma$ (MHz)					
$\delta(4\text{-}4)$	-6.4E-4±8.6E-4	-2.6E-4±6.0E-4	-0.0036±0.014	-2.7E-4±1.0E-4	7.2E-5±1.4E-4
$\delta(4\text{-}5)$	-5.8E-4±9.8E-4	-2.6E-4±7.9E-4	-0.0035±0.017	-2.6E-4±1.3E-4	7.9E-5±1.9E-4
$\delta(5\text{-}5)$	-7.6E-5±1.6E-4	-3.6E-6±5.5E-5	0.00037±0.0012	-3.0E-5±1.1E-5	2.6E-6±1.4E-5
$\delta(5\text{-}6)$	-6.9E-5±1.0E-4	-6.8E-6±7.6E-5	0.00038±0.0016	-2.7E-5±1.8E-5	4.3E-6±1.8E-5
Relative diff %					
$\delta(4\text{-}4)\%$	-0.021±0.028	0.22±0.5	<b>10.5±43.9</b>	-4.1±1.6	4.2±8.5
$\delta(5\text{-}6)\%$	<b>-0.0023±0.0035</b>	<b>0.0093±0.067</b>	<b>-1.1±4.6</b>	<b>-0.41±0.27</b>	<b>0.25±1.1</b>

#### 4. Conclusion: We need “Hot” MW spectra

This study was motivated by several interests: (1) our ambition to extend the 0.01-0.02 MHz (for  $A/B/C$ ) and 5-10 MHz (for  $J/K_a < 10\text{-}20$  transitions) ***prediction*** accuracy from the  $\text{SO}_2$  ground state MW spectra to vibrationally excited IR spectra, so that we can provide the most accurate hotband MW predictions for temperatures as high as 1000 K; (2) curiosity regarding the consistency of experimental data and current limits of the BTRHE strategy, i.e. to find out how much information we may extract from experimental IR/MW data, and how far we can further refine Ames prediction with it; (3) curiosity about the major limiting factors in future BTRHE applications, and the possibility to circumvent or mitigate them. Using the available high-resolution rovibrational EH(Expt) models and a complete set of BTRHE-based Ames IR line lists for  $\text{SO}_2$  isotopologues, we partially fulfilled these motivations by carrying out a systematic study on the  $A/B/C$  and quartic terms of EH(Expt) and EH(Ames).

Currently, our major observations include:

- (1) The EH(Ames) constants –  $A/B/C$  and quartic terms fitted from Ames  $\text{SO}_2$   $J=0\text{-}20$  energy levels – have recognizable and consistent patterns along vibrational quanta and wavenumber. The EH(Expt) constants have similar patterns too, but the  $\delta(\text{Ames} - \text{Expt})$  differences show they have more numerical noise and less consistency.
- (2) The  $1\text{E}-4 \sim 1\text{E}-3 \text{ cm}^{-1}$  accuracy of experimental rovibrational IR line positions is a major factor affecting the EH(Expt) consistency across isotopologues and vibrational quanta. Our independent test has found that, compared to parameters fitted with  $1\text{E}-7 \text{ cm}^{-1}$  precision, the  $1\text{E}-4 \text{ cm}^{-1}$  truncation leads to  $A/B/C$  deviations as large as  $\pm 0.05\text{--}0.10$  MHz, and  $\pm 5\text{--}50\%$  relative deviations on  $\Delta_J$ ,  $\delta$  and  $\delta_K$ . Precision of  $1\text{E}-5 \text{ cm}^{-1}$  is required to reduce the  $\delta\%$  of  $\Delta_J$  and  $\delta$  down to  $\sim 1\%$ , and  $1\text{E}-6 \text{ cm}^{-1}$  precision if needed for  $\delta_K$ .

Many experimental EH analyses do not have sufficient data precision to determine the full set of EH(Expt) parameters for one or more specific bands.

- (3) To improve the accuracy and consistency of EH(Expt) parameters, the only solution we can propose is to take advantage of the high accuracy of microwave or sub-millimeter spectra. Ideally, EH(Expt) analysis should be carried out with both “Hot” MW and IR data. We sincerely hope all future EH(Expt) analyses are at least partially based on MW data that carry 1E-6/1E-7 cm<sup>-1</sup> (or 3-30 kHz) or higher precision. Ulenikov *et al.*[35] has provided such an example where the  $\sigma_{\text{RMS}}$  has achieved the resolution of both IR and MW measurements.
- (4) “Hot” MW spectra are important to future molecular rovibrational spectral studies. It may become a synergy point of experiment, model, and theory. Experimental “hot” MW analysis will help purify existing EH(Expt) models, and provide more reliable constraints on future BTRHE line list projects. Running EH analysis on BTRHE IR line lists and finding the pattern of EH parameters may facilitate the hotband MW analysis, including predictions of the unperturbed EH parameters and reasonable estimates for the necessary coupling terms. One may utilize the A/B/C and quartic constant patterns to try assigning the vibrational states associated with the hot bands in MW or sub-MM spectra.
- (5) In general, we recommend filling in the highly accurate EH data along the ladder of vibrational state energies. The SO<sub>2</sub> energy ladder is 2v<sub>2</sub> → v<sub>1</sub> → v<sub>3</sub> → 3v<sub>2</sub> → v<sub>1</sub>+v<sub>2</sub> → v<sub>2</sub>+v<sub>3</sub> → ...etc. Note this ladder does not suggest measuring difference or sequence bands. Highly perturbed transitions are expected to provide more insights for the interactions among the states. Such data would significantly improve the coverage of experimental EH models. Our BTRHE refinements would take the absolute energies of perturbed levels, or their difference as reliable input. To improve the BTRHE predictions for SO<sub>2</sub> minor isotopologues, we suggest that hot-MW analysis should be done first for the lowest vibrational states of 626 → 636 → 646 → 666 and 626 → 627 → 628/727 → 728 → 828.
- (6) Some EH(Expt) models included the lowest order of Coriolis coupling term C<sup>1</sup>. We have carried out a case study on the 626 v<sub>1</sub>/v<sub>3</sub>/2v<sub>2</sub> triad. All the EH(Expt) constants are C<sup>1</sup> dependent, including A/B/C. They should *not* be compared with our EH(Ames) and the EH(Expt) without C<sup>1</sup>. See the case study details reported in the Supplementary files.
- (7) The consistency and systematic patterns of BTRHE based EH parameters may help identify unreliable EH(Expt) parameters, and provide reliable alternatives or predictions. Given accurate experimental EH parameters, for now we have two kinds of 0<sup>th</sup>-order refinements available:

$$X_{\text{predict}}^{v_1v_2v_3} = X_{\text{Ames}}^{v_1v_2v_3} - [X_{\text{Ames}}^{\text{GS}} - X_{\text{Expt}}^{\text{GS}}] = X_{\text{Expt}}^{\text{GS}} + X_{\text{Ames}}^{v_1v_2v_3} - X_{\text{Ames}}^{\text{GS}}, \text{ and}$$

$$X_{\text{predict}}^{\text{ISO2}} = X_{\text{Ames}}^{\text{ISO2}} - [X_{\text{Ames}}^{\text{ISO1}} - X_{\text{Expt}}^{\text{ISO1}}] = X_{\text{Expt}}^{\text{ISO1}} + X_{\text{Ames}}^{\text{ISO2}} - X_{\text{Ames}}^{\text{ISO1}}, \text{ where } X = A/B/C, \text{ or } \Delta_K, \Delta_{JK}, \dots$$

- (8) On the other hand, we did run least-squares “fit-on-fit” tests trying to formalize the patterns of EH(Ames) and / or EH(Expt) constants, but did not gain much except verifying that the EH(Ames) consistency is usually higher than that of EH(Expt) by one order of magnitude. It is not as easy as we had expected, because we do not have enough information to identify all the outliers, to assign proper weights for multiple EH(Expt) entries, and to separate the coupling perturbation effects from the data uncertainty (noise) from normal fitting errors of least-squares analysis. This is also true for the fits using the EH(Ames) parameters of lowest 40-80 vibrational states of SO<sub>2</sub> isotopologues: 626, 636, 646, 627, 628 and 828. Questions remain about how to *justify* the balance we choose between obtaining smaller  $\sigma_{\text{RMS}}$  and preventing overfits. We plan to revisit the topic in the future. Interested readers can contact us for numerical details.

There exist studies improving spectral data purity and EH(Expt) parameter reliability within the scope of current rovibrational IR resolution. For example, the MARVEL analysis on  $\text{SO}_2$  [60] tried to purify experimental line positions and energy levels, and their data products are at best self-consistent but are still limited by the 1E-3 or 1E-4  $\text{cm}^{-1}$  precision. More than 30 years ago, Ulenikov *et al* [61,62] reported mass-dependent formula that use the main isotopologue EH parameters to predict the EH parameters of minor isotopologues. These first require that the EH(Expt) be very accurately determined for the main isotopologue. The linear approximations used the EH parameters of lower quanta states to predict those of higher vibrational quanta states, which is similar to the  $\Delta C$  formula we give in Sec.3.1. However, we now know that the linear approximation only works in a limited range and gives low accuracy predictions. They also require EH(Expt) to be accurately determined for at least 2-3 lower energy states. New EH(Expt) analyses using hotband MW data will significantly improve the accuracy of those efforts, as it fixes these problems from the source. Its advantages can be maximized when combined with the consistency and patterns from high quality BTRHE line lists.

One potential application we have identified for EH(Ames) constants, including  $A/B/C$  and the quartic centrifugal distortion terms, is to provide reliable predictions for EH(Expt) analysis where existing experimental data does not support reliable determination for those terms. For example, both Ref.[32] and Ref.[58] adopted theoretically predicted higher order constants for the  $\text{SO}_2$  626  $2\nu_2$  state, because reasonable higher order terms are required for the  $\nu_1/2\nu_2$  resonance. Predictions were made on simple extrapolations from  $\text{GS} \rightarrow \nu_2$ , etc. We found such approximations can be safely replaced by EH(Ames) constants. The fitting  $\sigma_{\text{RMS}}$  is comparable or slightly better. This further supports our assertion regarding the EH(Ames) reliability. Interested readers can find more details in our case study report attached in supplementary files.

In addition to hot MW bands, hot Far-IR bands may also have merit for future study. Table 5 suggests the experimental EH models published for the  $\nu_2$  band of  $\text{SO}_2$  isotopologues have consistency much better than those of higher energy vibrational states. Accordingly, we expect  $\nu_2$ -related hot bands or any other strong enough hotbands around 500  $\text{cm}^{-1}$  can be measured with similar accuracy, e.g.  $210 \leftarrow 200$ , or  $2\nu_1 + \nu_2 \leftarrow 2\nu_1$ .

In the end, we re-emphasize two points:

- (1) the basic ideas and result are not new, but the current study should be a first of its kind to demonstrate quantitatively that the rovibrational EH(Expt) models need to improve their accuracy and consistency.
- (2) Obviously, the numerical results reported in this work are dependent on the specific BTRHE line lists and the specific  $A$ -reduced Watson Hamiltonian we used for the specific  $\text{SO}_2$  isotopologues, plus the experimental techniques and data accumulated, etc. The choice of  $A$ -reduction simply follows the experimental work in this field. Changing to  $S$ -reduction may affect the numerical results, but it is not expected to alter the observations and our conclusions. That is, unlike the vibrational ground states, the current experimental EH models for vibrationally excited states lack the consistency we need for the 2<sup>nd</sup>-order BTRHE implementation further refining our EH(Ames) parameters and line list predictions. Additionally, we are not really focusing on  $\text{SO}_2$ , it just happens to be the example we choose to discuss the topic. Of course, the choice of  $\text{SO}_2$  is simply because the Ames-296K line lists for 30  $\text{SO}_2$  isotopologues are available and suitable for this study.

To help readers verify our data and analysis, hundreds of data files are stored in a .tar supplementary file. In addition to the case study report, we attach the SPFIT files for the lowest 20-40 states of each isotopologue studied in this work, and files for the dozens of fits discussed in the case study report. We also include the original .xls spreadsheet files, ORIGIN project files, SPECTRO input/output, and a screenshot .pptx file for

readers having no access to the ORIGIN program. After downloading and decompressing the .tar file, readers please check the README.FIRST file.

## **Acknowledgement**

DWS, TJL, and XH gratefully acknowledge the financial support from the NASA 16-PDART16\_2-0 080, 17-APRA17-0051, and 18-APRA18-0013 grants. Huang acknowledges the NASA/SETI Co-operative Agreements NNX15AF45A, NNX17AL03G and 80NSSC19M0121. We sincerely thank our referees for their comments and suggestions which has greatly improved this paper, expanded our understanding, and helped to identifying the source of discrepancies between some EH(Expt) and EH(Ames). We also acknowledge stimulating discussions with Dr. Michael McCarthy (Harvard CfA) and Dr. John Pearson (JPL). The authors have known Dr. Jon Hougen and benefitted from stimulating scientific discussions with him since the 1980s. He brought an enthusiastic intellect to the discussions of molecular symmetry and was a true giant in such analyses. Dr. Hougen and his enthusiasm will be missed.

## References:

[1] J.J. Fortney, T.D. Robinson, S. Domagal-Goldman, D.S. Amundsen, M. Brogi, M. Claire, D. Crisp, E. Hebrard, H. Imanaka, R. de Kok, M.S. Marley, D. Teal, T. Barman, P. Bernath, A. Burrows, D. Charbonneau, R.S. Freedman, D. Gelino, C. Helling, K. Heng, A.G. Jensen, S. Kane, E.M.-R. Kempton, R.K. Kopparapu, N.K. Lewis, M. Lopez-Morales, J. Lyons, W. Lyra, V. Meadows, J. Moses, R. Pierrehumbert, O. Venot, S.X. Wang, J.T. Wright, The Need for Laboratory Work to Aid in The Understanding of Exoplanetary Atmospheres, (2016) 1–18. <http://arxiv.org/abs/1602.06305>.

[2] J.J. Fortney, T.D. Robinson, S. Domagal-Goldman, A.D. Del Genio, I.E. Gordon, E. Gharib-Nezhad, N. Lewis, C. Sousa-Silva, V. Airapetian, B. Drouin, R.J. Hargreaves, X. Huang, T. Karman, R.M. Ramirez, G.B. Rieker, J. Tennyson, R. Wordsworth, S.N. Yurchenko, A. V Johnson, T.J. Lee, C. Dong, S. Kane, M. Lopez-Morales, T. Fauchez, T. Lee, M.S. Marley, K. Sung, N. Haghighipour, T. Robinson, S. Horst, P. Gao, D. Kao, C. Dressing, R. Lupu, D.W. Savin, B. Fleury, O. Venot, D. Ascenzi, S. Milam, H. Linnartz, M. Gudipati, G. Gronoff, F. Salama, L. Gavilan, J. Bouwman, M. Turbet, Y. Benilan, B. Henderson, N. Batalha, R. Jensen-Clem, T. Lyons, R. Freedman, E. Schwertner, J. Goyal, L. Mancini, P. Irwin, J.-M. Desert, K. Molaverdikhani, J. Gizis, J. Taylor, J. Lothringer, R. Pierrehumbert, R. Zellem, N. Batalha, S. Rugheimer, J. Lustig-Yaeger, R. Hu, E. Kempton, G. Arney, M. Line, M. Alam, J. Moses, N. Iro, L. Kreidberg, J. Blecic, T. Louden, P. Molliere, K. Stevenson, M. Swain, K. Bott, N. Madhusudhan, J. Krissansen-Totton, D. Deming, I. Kitiashvili, E. Shkolnik, Z. Rustamkulov, L. Rogers, L. Close, The Need for Laboratory Measurements and Ab Initio Studies to Aid Understanding of Exoplanetary Atmospheres, (2019). <http://arxiv.org/abs/1905.07064>.

[3] X. Huang, D.W. Schwenke, T.J. Lee, Rovibrational spectra of ammonia. I. Unprecedented accuracy of a potential energy surface used with nonadiabatic corrections, *J. Chem. Phys.* 134 (2011) 044320. doi:10.1063/1.3541351.

[4] X. Huang, D.W. Schwenke, T.J. Lee, Rovibrational spectra of ammonia. II. Detailed analysis, comparison, and prediction of spectroscopic assignments for  $^{14}\text{NH}_3$ ,  $^{15}\text{NH}_3$ , and  $^{14}\text{ND}_3$ , *J. Chem. Phys.* 134 (2011) 044321. doi:10.1063/1.3541352.

[5] X. Huang, D.W. Schwenke, T.J. Lee, An accurate global potential energy surface, dipole moment surface, and rovibrational frequencies for  $\text{NH}_3$ , *J. Chem. Phys.* 129 (2008) 214304. doi:10.1063/1.3025885.

[6] K. Sung, L.R. Brown, X. Huang, D.W. Schwenke, T.J. Lee, S.L. Coy, K.K. Lehmann, Extended line positions, intensities, empirical lower state energies and quantum assignments of  $\text{NH}_3$  from 6300 to 7000  $\text{cm}^{-1}$ , *J. Quant. Spectrosc. Radiat. Transf.* 113 (2012) 1066–1083. doi:10.1016/j.jqsrt.2012.02.037.

[7] X. Huang, D.W. Schwenke, S.A. Tashkun, T.J. Lee, An isotopic-independent highly accurate potential energy surface for  $\text{CO}_2$  isotopologues and an initial  $^{12}\text{C}^{16}\text{O}_2$  infrared line list, *J. Chem. Phys.* 136 (2012) 124311. doi:10.1063/1.3697540.

[8] X. Huang (黃新川), D.W. Schwenke, R.S. Freedman, T.J. Lee, *Ames-2016 line lists for 13 isotopologues of  $\text{CO}_2$ : Updates, consistency, and remaining issues*, *J. Quant. Spectrosc. Radiat. Transf.* 203 (2017) 224–241. doi:10.1016/j.jqsrt.2017.04.026.

[9] X. Huang, D.W. Schwenke, T.J. Lee, Highly accurate potential energy surface, dipole moment surface, rovibrational energy levels, and infrared line list for  $^{32}\text{S}^{16}\text{O}_2$  up to 8000  $\text{cm}^{-1}$ , *J. Chem. Phys.* 140 (2014) 114311. doi:10.1063/1.4868327.

[10] D.S. Underwood, J. Tennyson, S.N. Yurchenko, X. Huang, D.W. Schwenke, T.J. Lee, S. Clausen, A. Fateev, *ExoMol molecular line lists - XIV. The rotation-vibration spectrum of hot  $\text{SO}_2$* , *Mon. Not. R. Astron. Soc.* 459 (2016) 3890–3899. doi:10.1093/mnras/stw849.

[11] X. Huang, D.W. Schwenke, T.J. Lee, *Empirical infrared line lists for five  $\text{SO}_2$  isotopologues:  $^{32/33/34/36}\text{S}^{16}\text{O}_2$  and  $^{32}\text{S}^{18}\text{O}_2$* , *J. Mol. Spectrosc.* 311 (2015) 19–24. doi:10.1016/j.jms.2015.01.010.

[12] X. Huang, D.W. Schwenke, T.J. Lee, Ames 32S16O18O line list for high-resolution experimental IR analysis, *J. Mol. Spectrosc.* 330 (2016) 101–111. doi:10.1016/j.jms.2016.08.013.

[13] X. Huang, D.W. Schwenke, T.J. Lee, Quantitative validation of Ames IR intensity and new line lists for 32/33/34S16O2, 32S18O2 and 16O32S18O, *J. Quant. Spectrosc. Radiat. Transf.* 225 (2019) 327–336. doi:10.1016/j.jqsrt.2018.11.039.

[14] X. Huang, D.W. Schwenke, T.J. Lee, Isotopologue consistency of semi-empirically computed infrared line lists and further improvement for rare isotopologues: CO2 and SO2 case studies, *J. Quant. Spectrosc. Radiat. Transf.* 230 (2019) 222–246. doi:10.1016/j.jqsrt.2019.03.002.

[15] I.E. Gordon, L.S. Rothman, C. Hill, R.V. Kochanov, Y. Tan, P.F. Bernath, M. Birk, V. Boudon, A. Campargue, K.V. Chance, B.J. Drouin, J.-M. Flaud, R.R. Gamache, J.T. Hodges, D. Jacquemart, V.I. Perevalov, A. Perrin, K.P. Shine, M.-A.H. Smith, J. Tennyson, G.C. Toon, H. Tran, V.G. Tyuterev, A. Barbe, A.G. Császár, V.M. Devi, T. Furtenbacher, J.J. Harrison, J.-M. Hartmann, A. Jolly, T.J. Johnson, T. Karman, I. Kleiner, A.A. Kyuberis, J. Loos, O.M. Lyulin, S.T. Massie, S.N. Mikhailenko, N. Moazzen-Ahmadi, H.S.P. Müller, O.V. Naumenko, A.V. Nikitin, O.L. Polyansky, M. Rey, M. Rotger, S.W. Sharpe, K. Sung, E. Starikova, S.A. Tashkun, J. Vander Auwera, G. Wagner, J. Wilzewski, P. Wcisło, S. Yu, E.J. Zak, *The HITRAN2016 molecular spectroscopic database*, *J. Quant. Spectrosc. Radiat. Transf.* 203 (2017) 3–69. doi:10.1016/j.jqsrt.2017.06.038.

[16] L.S. Rothman, I.E. Gordon, Y. Babikov, A. Barbe, D. Chris Benner, P.F. Bernath, M. Birk, L. Bizzocchi, V. Boudon, L.R. Brown, A. Campargue, K. Chance, E.A. Cohen, L.H. Coudert, V.M. Devi, B.J. Drouin, A. Fayt, J.-M. Flaud, R.R. Gamache, J.J. Harrison, J.-M. Hartmann, C. Hill, J.T. Hodges, D. Jacquemart, A. Jolly, J. Lamouroux, R.J. Le Roy, G. Li, D.A. Long, O.M. Lyulin, C.J. Mackie, S.T. Massie, S. Mikhailenko, H.S.P. Müller, O.V. Naumenko, A.V. Nikitin, J. Orphal, V. Perevalov, A. Perrin, E.R. Polovtseva, C. Richard, M.A.H. Smith, E. Starikova, K. Sung, S. Tashkun, J. Tennyson, G.C. Toon, V.G. Tyuterev, G. Wagner, *The HITRAN2012 molecular spectroscopic database*, *J. Quant. Spectrosc. Radiat. Transf.* 130 (2013) 4–50. doi:10.1016/j.jqsrt.2013.07.002.

[17] L.S.S. Rothman, I.E.E. Gordon, A. Barbe, D.C.C. Benner, P.F.F. Bernath, M. Birk, V. Boudon, L.R.R. Brown, A. Campargue, J.-P.P. Champion, K. Chance, L.H.H. Coudert, V. Dana, V.M.M. Devi, S. Fally, J.-M.M. Flaud, R.R.R. Gamache, A. Goldman, D. Jacquemart, I. Kleiner, N. Lacome, W.J.J. Lafferty, J.-Y.Y. Mandin, S.T.T. Massie, S.N.N. Mikhailenko, C.E.E. Miller, N. Moazzen-Ahmadi, O.V. V. Naumenko, A.V. V. Nikitin, J. Orphal, V.I.I. Perevalov, A. Perrin, A. Predoi-Cross, C.P.P. Rinsland, M. Rotger, M. Šimečková, M.A.H.A.H. Smith, K. Sung, S.A.A. Tashkun, J. Tennyson, R.A.A. Toth, A.C.C. Vandaele, J. Vander Auwera, The HITRAN 2008 molecular spectroscopic database, *J. Quant. Spectrosc. Radiat. Transf.* 110 (2009) 533–572. doi:10.1016/j.jqsrt.2009.02.013.

[18] H.S.P. Müller, F. Schlöder, J. Stutzki, G. Winnewisser, *The Cologne Database for Molecular Spectroscopy, CDMS: a useful tool for astronomers and spectroscopists*, *J. Mol. Struct.* 742 (2005) 215–227. doi:10.1016/j.molstruc.2005.01.027.

[19] C.P. Endres, S. Schlemmer, P. Schilke, J. Stutzki, H.S.P. Müller, The Cologne Database for Molecular Spectroscopy, CDMS, in the Virtual Atomic and Molecular Data Centre, VAMDC, *J. Mol. Spectrosc.* 327 (2016) 95–104. doi:10.1016/j.jms.2016.03.005. See more at <https://cdms.astro.uni-koeln.de>.

[20] W.J. Lafferty, G.T. Fraser, A.S. Pine, J.-M. Flaud, C. Camy-Peyret, V. Dana, J.-Y. Mandin, A. Barbe, J.J. Plateaux, S. Bouazza, The 3v3 band of 32S16O2: Line positions and intensities, *J. Mol. Spectrosc.* 154 (1992) 51–60. doi:10.1016/0022-2852(92)90028-M.

[21] J.M. Flaud, W.J. Lafferty, 32S16O2: A Refined Analysis of the 3v3 Band and Determination of Equilibrium Rotational Constants, *J. Mol. Spectrosc.* 161 (1993) 396–402. doi:10.1006/jmsp.1993.1245.

[22] W.J. Lafferty, A.S. Pine, G. Hilpert, R.L. Sams, J.-M. Flaud, The v1+ v3 and 2v1+ v3 Band Systems of SO2: Line Positions and Intensities, *J. Mol. Spectrosc.* 176 (1996) 280–286. doi:10.1006/jmsp.1996.0088.

[23] W.J. Lafferty, J.-M. Flaud, G. Guelachvili, Analysis of the 2v1 Band System of SO2, *J. Mol. Spectrosc.* 188 (1998) 106–107. doi:10.1006/jmsp.1997.7493.

[24] S. Belov, M. Tretyakov, I. Kozin, E. Klisch, G. Winnewisser, W. Lafferty, J.-M. Flaud, High Frequency Transitions in the Rotational Spectrum of SO2, *J. Mol. Spectrosc.* 191 (1998) 17–27. doi:10.1006/jmsp.1998.7576.

[25] H.S.P. Müller, J. Farhoomand, E.A. Cohen, B. Brupbacher-Gatehouse, M. Schafer, A. Bauder, G. Winnewisser, The

rotational spectrum of SO<sub>2</sub> and the determination of the hyperfine constants and nuclear magnetic shielding tensors of (SO<sub>2</sub>)-S-33 and (SOO)-O-17, *J. Mol. Spectrosc.* 201 (2000) 1–8. doi:10.1006/jmsp.2000.8072.

[26] H.S.P. Müller, S. Brünken, Accurate rotational spectroscopy of sulfur dioxide, SO<sub>2</sub>, in its ground vibrational and first excited bending states,  $v_2=0, 1$ , up to 2 THz, *J. Mol. Spectrosc.* 232 (2005) 213–222. doi:10.1016/j.jms.2005.04.010.

[27] O.N. Ulenikov, E.S. Bekhtereva, S. Alanko, V.M. Horneman, O.V. Gromova, C. Leroy, *On the high resolution spectroscopy and intramolecular potential function of SO<sub>2</sub>*, *J. Mol. Spectrosc.* 257 (2009) 137–156. doi:10.1016/j.jms.2009.07.005.

[28] O.N. Ulenikov, E.S. Bekhtereva, V.-M. Horneman, S. Alanko, O.V. Gromova, *High resolution study of the 3v<sub>1</sub> band of SO<sub>2</sub>*, *J. Mol. Spectrosc.* 255 (2009) 111–121. doi:10.1016/j.jms.2009.03.009.

[29] O.N. Ulenikov, E.S. Bekhtereva, O.V. Gromova, S. Alanko, V.-M. Horneman, C. Leroy, *Analysis of highly excited ‘hot’ bands in the SO<sub>2</sub> molecule: v<sub>2</sub>+3v<sub>3</sub>-v<sub>2</sub> and 2v<sub>1</sub>+v<sub>2</sub>+v<sub>3</sub>-v<sub>2</sub>*, *Mol. Phys.* 108 (2010) 1253–1261. doi:10.1080/00268970903468297.

[30] O.N. Ulenikov, O.V. Gromova, E.S. Bekhtereva, I.B. Bolotova, C. Leroy, V.-M. Horneman, S. Alanko, *High resolution study of the and “hot” bands and ro-vibrational re-analysis of the polyad of the <sup>32</sup>SO<sub>2</sub> molecule*, *J. Quant. Spectrosc. Radiat. Transf.* 112 (2011) 486–512. doi:10.1016/j.jqsrt.2010.09.013.

[31] O.N. Ulenikov, O.V. Gromova, E.S. Bekhtereva, I.B. Bolotova, I.A. Konov, V.-M. Horneman, C. Leroy, *High resolution analysis of the SO<sub>2</sub> spectrum in the 2600–2900 cm<sup>-1</sup> region: 2v<sub>3</sub>, v<sub>2</sub>+2v<sub>3</sub>-v<sub>2</sub> and 2v<sub>1</sub>+v<sub>2</sub> bands*, *J. Quant. Spectrosc. Radiat. Transf.* 113 (2012) 500–517. doi:10.1016/j.jqsrt.2012.01.006.

[32] O.N. Ulenikov, G.A. Onopenko, O.V. Gromova, E.S. Bekhtereva, V.-M. Horneman, Re-analysis of the (100), (001), and (020) rotational structure of SO<sub>2</sub> on the basis of high resolution FTIR spectra, *J. Quant. Spectrosc. Radiat. Transf.* 130 (2013) 220–232. doi:10.1016/j.jqsrt.2013.04.011.

[33] O.N. Ulenikov, O.V. Gromova, E.S. Bekhtereva, A.S. Belova, S. Bauerecker, C. Maul, C. Sydow, V.-M. Horneman, *High resolution analysis of the (111) vibrational state of SO<sub>2</sub>*, *J. Quant. Spectrosc. Radiat. Transf.* 144 (2014) 1–10. doi:10.1016/j.jqsrt.2014.03.027.

[34] O.N. Ulenikov, E.S. Bekhtereva, O.V. Gromova, K.B. Berezkin, V.M. Horneman, C. Sydow, C. Maul, S. Bauerecker, *First high resolution analysis of the 3v<sub>2</sub> and 3v<sub>2</sub>-v<sub>2</sub> bands of <sup>32</sup>S<sup>16</sup>O<sub>2</sub>*, *J. Quant. Spectrosc. Radiat. Transf.* 202 (2017) 1–5. doi:10.1016/j.jqsrt.2017.07.012.

[35] O.N. Ulenikov, E.S. Bekhtereva, O.V. Gromova, M. Quack, G.C. Mellau, C. Sydow, S. Bauerecker, *Extended analysis of the high resolution FTIR spectrum of <sup>32</sup>S<sup>16</sup>O<sub>2</sub> in the region of the v<sub>2</sub> band: Line positions, strengths, and pressure broadening widths*, *J. Quant. Spectrosc. Radiat. Transf.* 210 (2018) 141–155. doi:10.1016/j.jqsrt.2018.02.010.

[36] E. Klisch, P. Schilke, S. Belov, G. Winnewisser, *33SO<sub>2</sub>: Interstellar Identification and Laboratory Measurements*, *J. Mol. Spectrosc.* 186 (1997) 314–8. <http://www.ncbi.nlm.nih.gov/pubmed/9446769>.

[37] J.-M. Flaud, T.A. Blake, W.J. Lafferty, *First high-resolution analysis of the v<sub>1</sub>, v<sub>3</sub> and v<sub>1</sub>+v<sub>3</sub> bands of sulphur dioxide <sup>33</sup>S<sup>16</sup>O<sub>2</sub>*, *Mol. Phys.* 115 (2017) 447–453. doi:10.1080/00268976.2016.1269966.

[38] T.A. Blake, J.-M. Flaud, W.J. Lafferty, *First analysis of the rotationally-resolved v<sub>2</sub> and 2v<sub>2</sub>-v<sub>2</sub> bands of sulfur dioxide, <sup>33</sup>S<sup>16</sup>O<sub>2</sub>*, *J. Mol. Spectrosc.* 333 (2017) 19–22. doi:10.1016/j.jms.2016.12.011.

[39] W.J. Lafferty, J.-M. Flaud, R.L. Sams, E.H. Abib Ngom, High resolution analysis of the rotational levels of the (000), (010), (100), (001), (020), (110) and (011) vibrational states of 34S16O2, *J. Mol. Spectrosc.* 252 (2008) 72–76. doi:10.1016/j.jms.2008.06.013.

[40] W.J. Lafferty, J.-M. Flaud, E.H.A. Ngom, R.L. Sams, *34S16O2: High-resolution analysis of the (030), (101), (111), (002) and (201) vibrational states; determination of equilibrium rotational constants for sulfur dioxide and anharmonic vibrational constants*, *J. Mol. Spectrosc.* 253 (2009) 51–54. doi:10.1016/j.jms.2008.09.006.

[41] J.M. Flaud, W.J. Lafferty, R.L. Sams, Line intensities for the  $v_1$ ,  $v_3$  and  $v_1+v_3$  bands of  $^{34}\text{SO}_2$ , *J. Quant. Spectrosc. Radiat. Transf.* 110 (2009) 669–674. doi:10.1016/j.jqsrt.2008.12.003.

[42] O.N. Ulenikov, E.S. Bekhtereva, O.V. Gromova, V.A. Zamotaeva, E.A. Sklyarova, C. Sydow, C. Maul, S. Bauerecker, *First high resolution analysis of the  $2v_1$ ,  $2v_3$ , and  $v_1+v_3$  bands of  $^{32}\text{S}^{18}\text{O}_2$* , *J. Quant. Spectrosc. Radiat. Transf.* 185 (2016) 12–21. doi:10.1016/j.jqsrt.2016.08.008.

[43] O.N. Ulenikov, E.S. Bekhtereva, O.V. Gromova, T. Buttersack, C. Sydow, S. Bauerecker, *High resolution FTIR study of  $^{34}\text{S}^{16}\text{O}_2$ : Re-analysis of the bands  $v_1+v_2, v_2+v_3$ , and first analysis of the hot band  $2v_2+v_3-v_2$* , *J. Mol. Spectrosc.* 319 (2016) 17–25. doi:10.1016/j.jms.2015.11.003.

[44] F. Gueye, L. Manceron, A. Perrin, F.K. Tchana, J. Demaison, First far-infrared high-resolution analysis of the  $v_2$  band of sulphur dioxide  $^{32}\text{S}^{16}\text{O}^{18}\text{O}$  and  $^{32}\text{S}^{18}\text{O}^2$ , *Mol. Phys.* 114 (2016) 2769–2776. doi:10.1080/00268976.2016.1154619.

[45] O.N. Ulenikov, E.S. Bekhtereva, Y.V. Krivchikova, V.A. Zamotaeva, T. Buttersack, C. Sydow, S. Bauerecker, *<math>\langle i \rangle</math>Study of the high resolution spectrum of  $^{32}\text{S}^{16}\text{O}^{18}\text{O}$ : The  $v_1$  and  $v_3$  bands<math>\langle i \rangle</math>*, *J. Quant. Spectrosc. Radiat. Transf.* 168 (2016) 29–39. doi:10.1016/j.jqsrt.2015.08.010.

[46] O.N. Ulenikov, E.S. Bekhtereva, O.V. Gromova, V.A. Zamotaeva, S.I. Kuznetsov, C. Sydow, S. Bauerecker, *High resolution study of the rotational structure of doubly excited vibrational states of  $^{32}\text{S}^{16}\text{O}^{18}\text{O}$ : The first analysis of the  $2v_1$ ,  $v_1+v_3$ , and  $2v_3$  bands*, *J. Quant. Spectrosc. Radiat. Transf.* 189 (2017) 344–350. doi:10.1016/j.jqsrt.2016.12.019.

[47] O.N. Ulenikov, E.S. Bekhtereva, O.V. Gromova, V.M. Horneman, C. Sydow, S. Bauerecker, *<math>\langle i \rangle</math>High resolution FTIR spectroscopy of sulfur dioxide in the  $1550\text{--}1950\text{ cm}^{-1}$  region: First analysis of the  $v_1+v_2/v_2+v_3$  bands of  $^{32}\text{S}^{16}\text{O}^{18}\text{O}$  and experimental line intensi*, *J. Quant. Spectrosc. Radiat. Transf.* 203 (2017) 377–391. doi:10.1016/j.jqsrt.2017.02.005.

[48] O.N. Ulenikov, O.V. Gromova, E.S. Bekhtereva, A.G. Ziatkova, E.A. Sklyarova, S.I. Kuznetsov, C. Sydow, S. Bauerecker, *First rotational analysis of the (111) and (021) vibrational state of  $\text{S}^{16}\text{O}^{18}\text{O}$  from the “hot”  $v_1+v_2+v_3-v_2$  and  $2v_2+v_3-v_2$  bands*, *J. Quant. Spectrosc. Radiat. Transf.* 202 (2017) 98–103. doi:10.1016/j.jqsrt.2017.07.029.

[49] van R. R, Steenbeckeliers G. Spectre de rotattion de la molecule  $^{32}\text{S}^{16}\text{O}^{18}\text{O}$ ., *Ann Soc Sc Brux.* 97 (1984) 117–153.

[50] O.N. Ulenikov, E.S. Bekhtereva, Y.V. Krivchikova, Y.B. Morzhikova, T. Buttersack, C. Sydow, S. Bauerecker, *High resolution analysis of  $^{32}\text{S}^{18}\text{O}_2$  spectra: The  $v_1$  and  $v_3$  interacting bands*, *J. Quant. Spectrosc. Radiat. Transf.* 166 (2015) 13–22. doi:10.1016/j.jqsrt.2015.07.004.

[51] O.N. Ulenikov, E.S. Bekhtereva, O.V. Gromova, V.A. Zamotaeva, S.I. Kuznetsov, C. Sydow, C. Maul, S. Bauerecker, *First high resolution analysis of the  $v_1+v_2$  and  $v_2+v_3$  bands of  $^{32}\text{S}^{18}\text{O}_2$* , *J. Quant. Spectrosc. Radiat. Transf.* 179 (2016) 187–197. doi:10.1016/j.jqsrt.2016.03.038.

[52] O.N. Ulenikov, O.V. Gromova, E.S. Bekhtereva, Y.B. Morzhikova, C. Maul, C. Sydow, S. Bauerecker, *Study of highly excited ro-vibrational states of  $\text{S}^{18}\text{O}^2$  from “hot” transitions: The bands  $v_1+v_2+v_3-v_2$ ,  $2v_1+v_2-v_2$ , and  $2v_2+v_3-v_2$* , *J. Quant. Spectrosc. Radiat. Transf.* 196 (2017) 159–164. doi:10.1016/j.jqsrt.2017.04.004.

[53] S.E. Novick, A beginner’s guide to Pickett’s SPCAT/SPFIT, *J. Mol. Spectrosc.* 329 (2016) 1–7. doi:10.1016/j.jms.2016.08.015.

[54] H.M. Pickett, *The fitting and prediction of vibration-rotation spectra with spin interactions*, *J. Mol. Spectrosc.* 148 (1991) 371–377. doi:10.1016/0022-2852(91)90393-O.

[55] H.S.P. Müller, Online EH(CDMS) model data : <https://cdms.astro.uni-koeln.de/classic/predictions/daten/SO2/>,  $^{34}\text{SO}_2/\text{s34.par}$  for 646;  $\text{SO}_2/\text{alt.1/so2.par}$  for 626;  $^{33}\text{SO}_2/\text{s33.par}$  for 636;  $\text{SO}^{18}\text{O}/\text{o18.par}$  for 628;  $\text{SO}^{17}\text{O}/\text{17.par}$  for 627, (n.d.).

[56] J.M. Dowling, Centrifugal distortion in planar molecules, *J. Mol. Spectrosc.* 6 (1961) 550–553. doi:10.1016/0022-

- [57] J.K.G. Watson, A planarity relation for sextic centrifugal distortion constants, *J. Mol. Spectrosc.* 65 (1977) 123–133. doi:10.1016/0022-2852(77)90364-2.
- [58] J.M. Flaud, A. Perrin, L.M. Salah, W.J. Lafferty, G. Guelachvili, A Reanalysis of the (010), (020), (100), and (001) Rotational Levels of  $32S16O_2$ , *J. Mol. Spectrosc.* 160 (1993) 272–278. doi:10.1006/jmsp.1993.1174.
- [59] O.N. Ulenikov, O.V. Gromova, E.S. Bekhtereva, Y.V. Krivchikova, E.A. Sklyarova, T. Buttersack, C. Sydow, S. Bauerecker, *High resolution FTIR study of  $^{34}S^{16}O_2$  : The bands  $2v_3$ ,  $2v_1+v_2$  and  $2v_1+v_2-v_2$* , *J. Mol. Spectrosc.* 318 (2015) 26–33. doi:10.1016/j.jms.2015.09.009.
- [60] R. Tóbiás, T. Furtenbacher, A.G. Császár, O. V. Naumenko, J. Tennyson, J.-M. Flaud, P. Kumar, B. Poirier, *Critical evaluation of measured rotational–vibrational transitions of four sulphur isotopologues of  $S^{16}O_2$* , *J. Quant. Spectrosc. Radiat. Transf.* 208 (2018) 152–163. doi:10.1016/j.jqsrt.2018.01.006.
- [61] A.D. Bykov, Y.S. Makushkin, O.N. Ulenikov, On isotope effects in polyatomic molecules, *J. Mol. Spectrosc.* 85 (1981) 462–479. doi:10.1016/0022-2852(81)90217-4.
- [62] A.D. Bykov, Y.S. Makushkin, O.N. Ulenikov, On the displacements of centers of vibration-rotation bands under isotope substitution in polyatomic molecules, *J. Mol. Spectrosc.* 93 (1982) 46–54. doi:10.1016/0022-2852(82)90273-9.

## \*Highlights (for review)

1. First systematic and quantitative study to compare the EH(Expt) and EH(BTRHE) models
2. The consistency of existing SO<sub>2</sub> EH(Expt) models is not high enough to guide the refinement of EH(BTRHE) and line lists, except for GS and nu\_2
3. The 1E-3 ~ 1E-4 cm<sup>-1</sup> data precision of most experimental rovibrational data severely hurts the EH(Expt) accuracy and consistency
4. "Hot" bands in the microwave and submillimeter region would be the most promising solution to improve the rovibrational IR analysis
5. An EH(Expt) analysis explicitly including the lowest order Coriolis coupling term should not be directly compared to other EH analyses excluding the term.

# Figure(s)

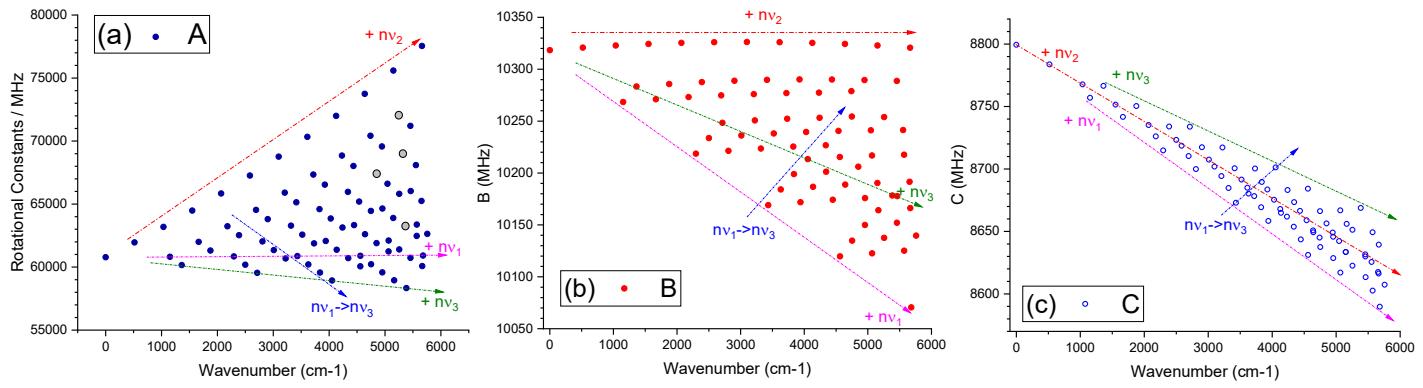


Fig.1 The vibrational dependence of A/B/C rotational constants of  $\text{SO}_2$  626, i.e.  ${}^{32}\text{S}{}^{16}\text{O}_2$ : a) A; b) B; c) C.

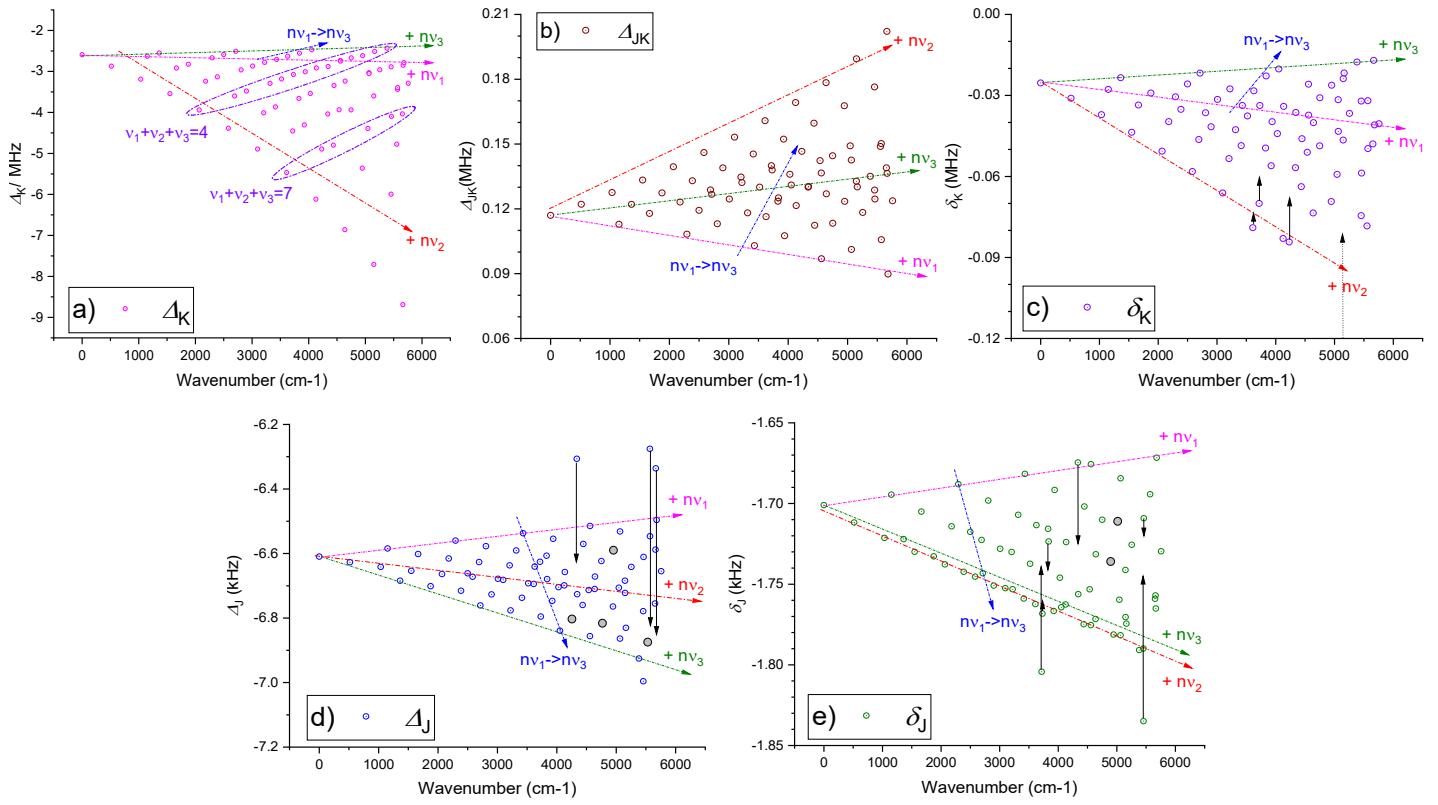


Fig.2 The vibrational dependence of Quartic centrifugal distortion constants of  $\text{SO}_2$  626, i.e.  ${}^{32}\text{S}{}^{16}\text{O}_2$ : a)  $\Delta_K$ ; b)  $\Delta_{JK}$ ; c)  $\delta_K$ ; d)  $\Delta_J$ ; e)  $\delta_J$

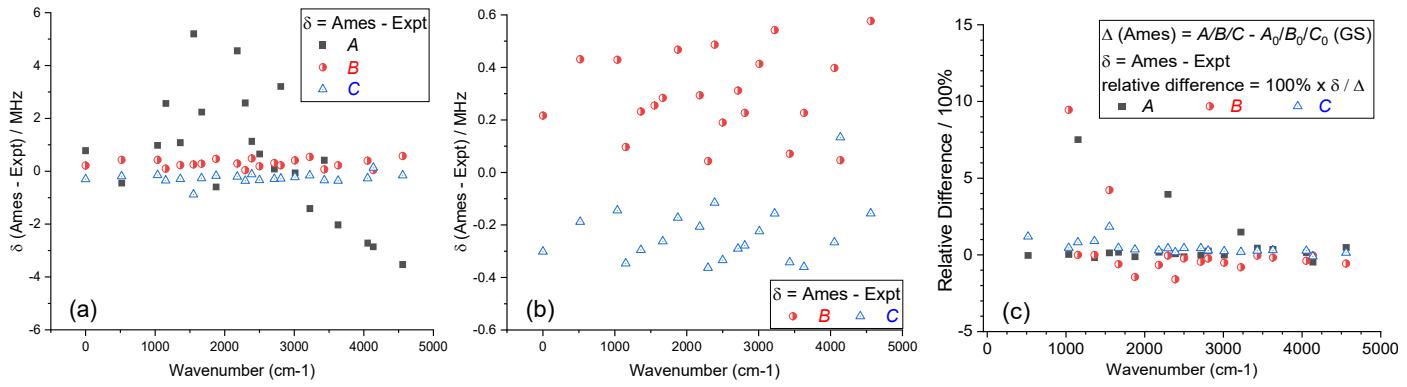


Fig.3 a) The  $\delta_{A/B/C}$  differences on 21  $\text{SO}_2$  626 vibrational states,  $\delta = X_{\text{Ames}} - X_{\text{Expt}}$ ; b) the  $\delta_B$  and  $\delta_C$  differences,  $\delta C(3\nu_2)$  is out of range ; c) the relative differences as defined in text and the plot.

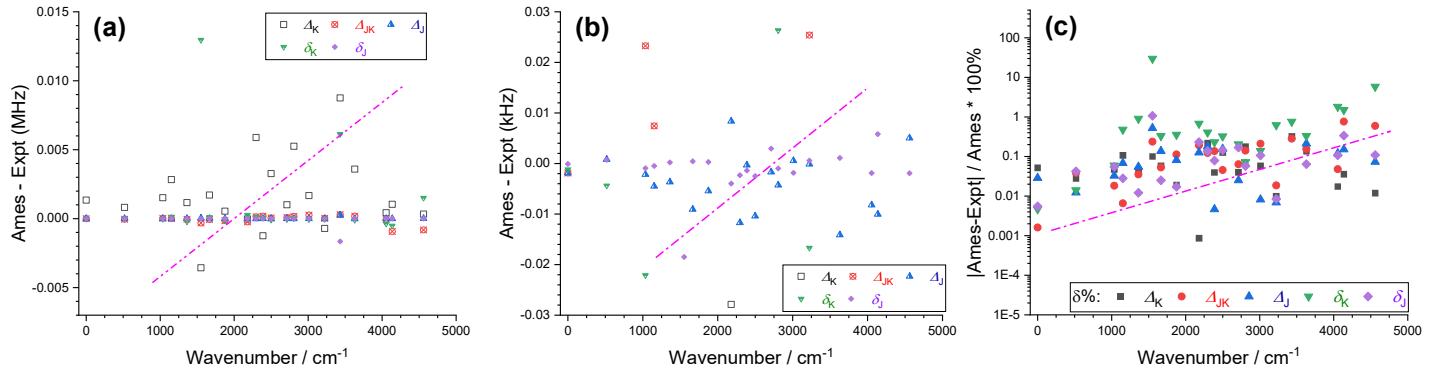


Fig.4 Ames vs Expt differences of five quartic centrifugal distortion constants of  $\text{SO}_2$  626 vibrational states: a)-b)  $\delta = X_{\text{Ames}} - X_{\text{Expt}}$ ; c) relative difference  $\delta\% = 100\% \times \delta/X_{\text{Ames}}$ .

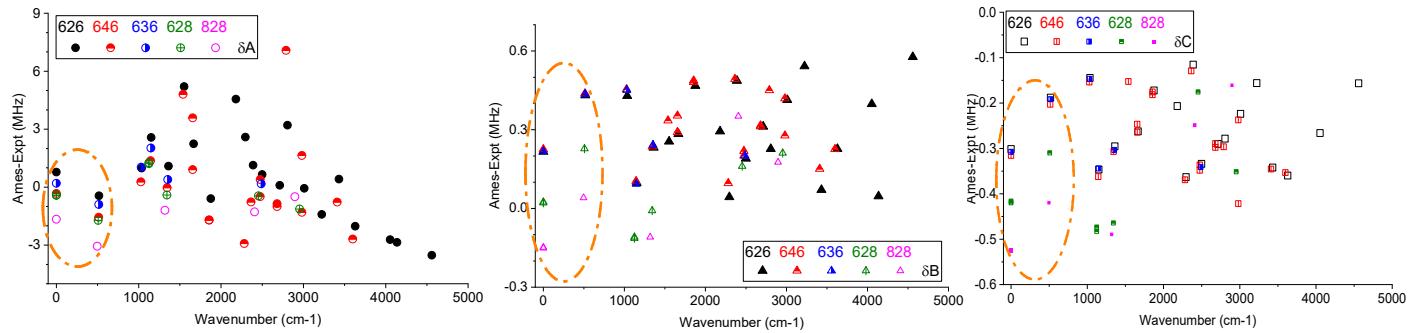


Fig.5 The  $\delta = X_{\text{Ames}} - X_{\text{Expt}}$  differences of A/B/C constants for 5 isotopologues: a) A; b) B; c) C.

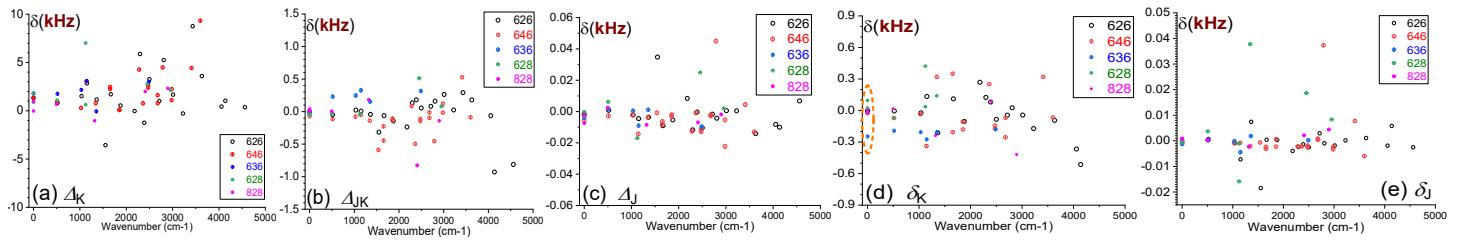


Fig.6  $\delta(X_{\text{Ames}}^{\text{ISO}} - X_{\text{Expt}}^{\text{ISO}})$  for the five quartic EH terms of 5  $\text{SO}_2$  isotopologues: a)  $\Delta_K$ ; b)  $\Delta_{JK}$ ; c)  $\Delta_J$ ; d)  $\delta_K$ ; e)  $\delta_J$ . in kHz.

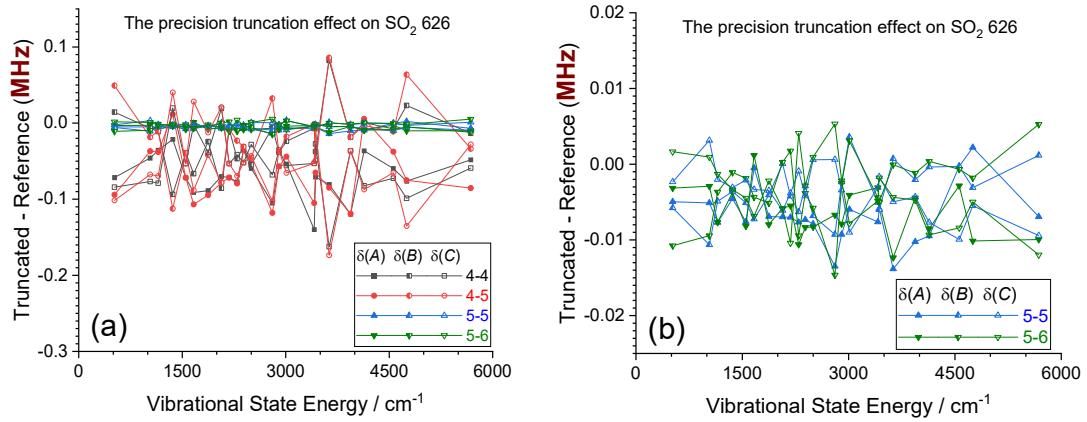


Fig.7 A/B/C differences between the EH(truncated) and EH(1E-7) reference values. a) four truncated sets in a range of 0.3 MHz; b) 1E-5  $\text{cm}^{-1}$  (0.3 MHz) truncation differences in a range of 0.02 MHz.

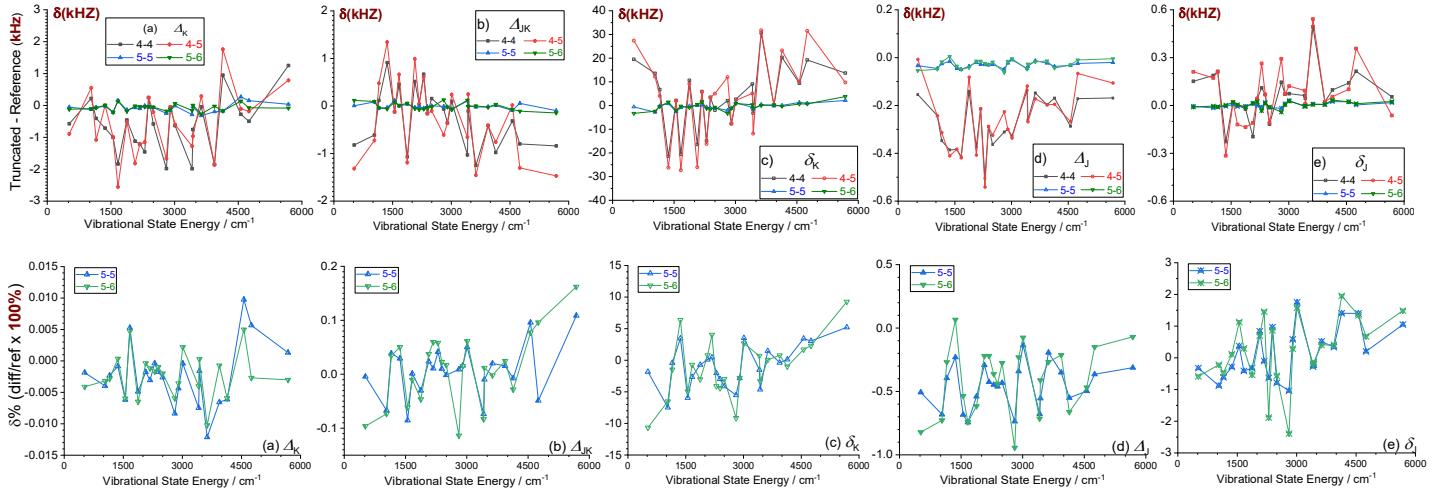


Fig.8  $\delta = X(\text{truncated}) - X(\text{reference 1E-7})$  differences (top row) and relative differences  $\delta\% = \delta / X(\text{reference 1E-7})$  (bottom row) for 5 quartic constants, from 4 truncation tests: a)  $\Delta_K$ ; b)  $\Delta_{JK}$ ; c)  $\delta_K$ ; d)  $\Delta_J$ ; e)  $\delta_J$

# Figure(s)

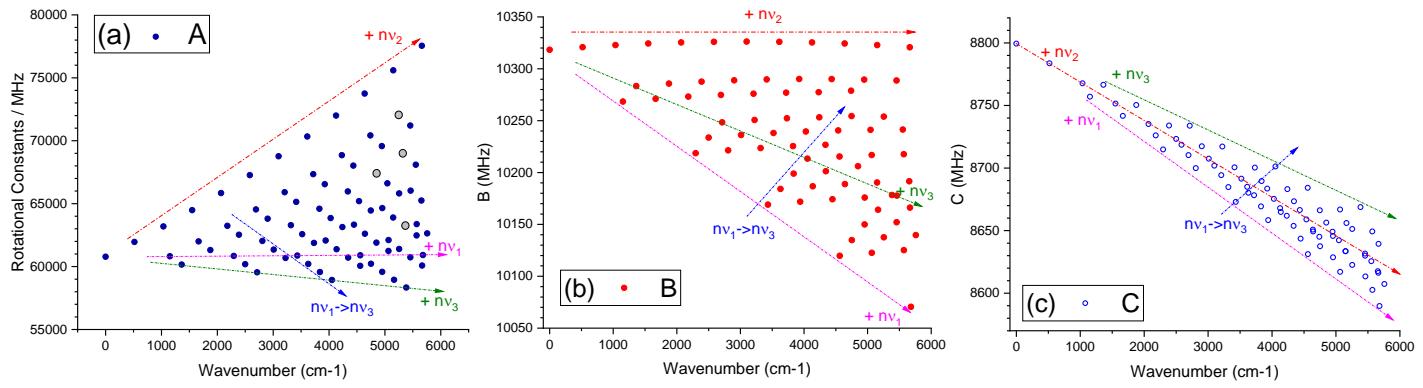


Fig.1 The vibrational dependence of A/B/C rotational constants of  $\text{SO}_2$  626, i.e.  $^{32}\text{S}^{16}\text{O}_2$ : a) A; b) B; c) C.

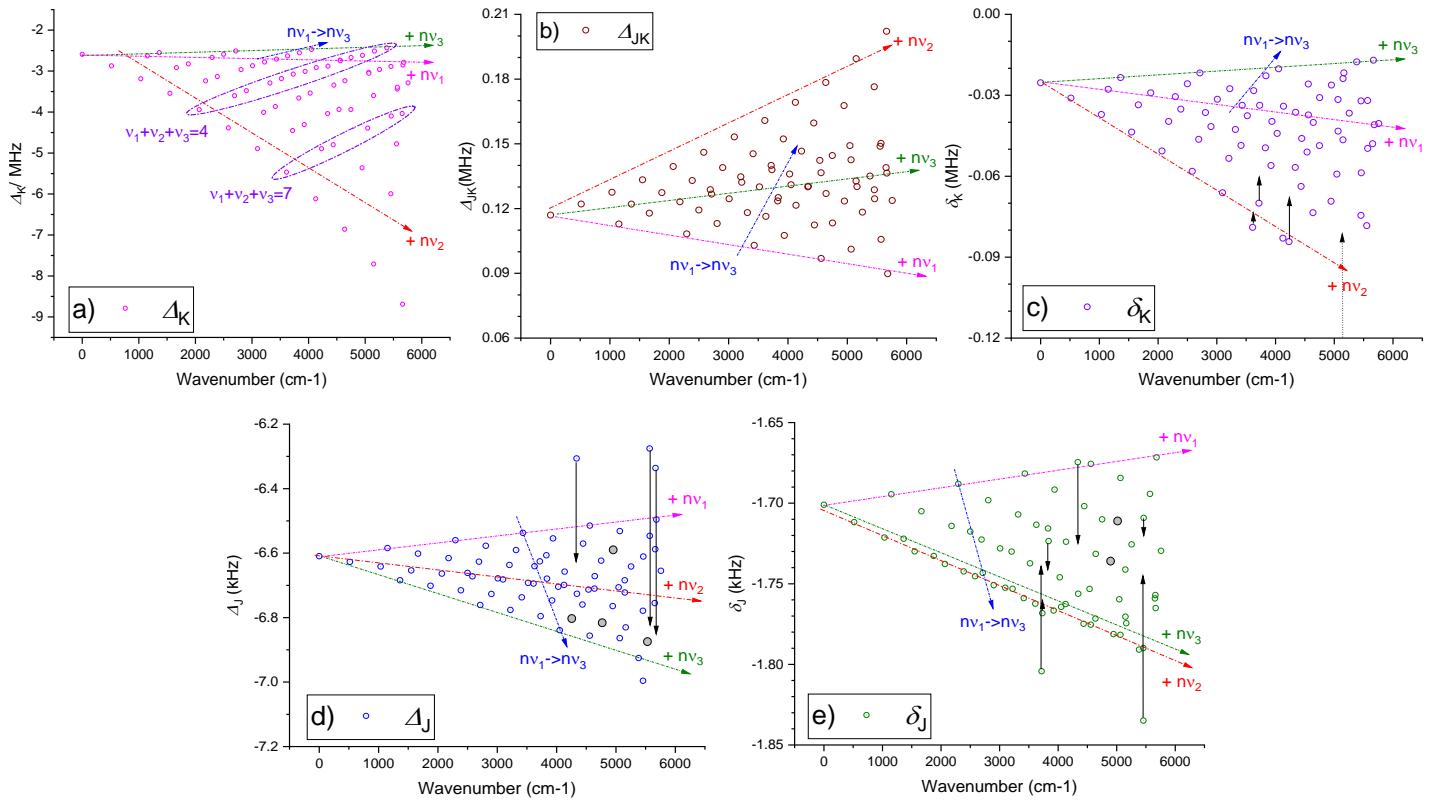


Fig.2 The vibrational dependence of Quartic centrifugal distortion constants of  $\text{SO}_2$  626, i.e.  $^{32}\text{S}^{16}\text{O}_2$ : a)  $\Delta_K$ ; b)  $\Delta_{JK}$ ; c)  $\delta_K$ ; d)  $\Delta_J$ ; e)  $\delta_J$

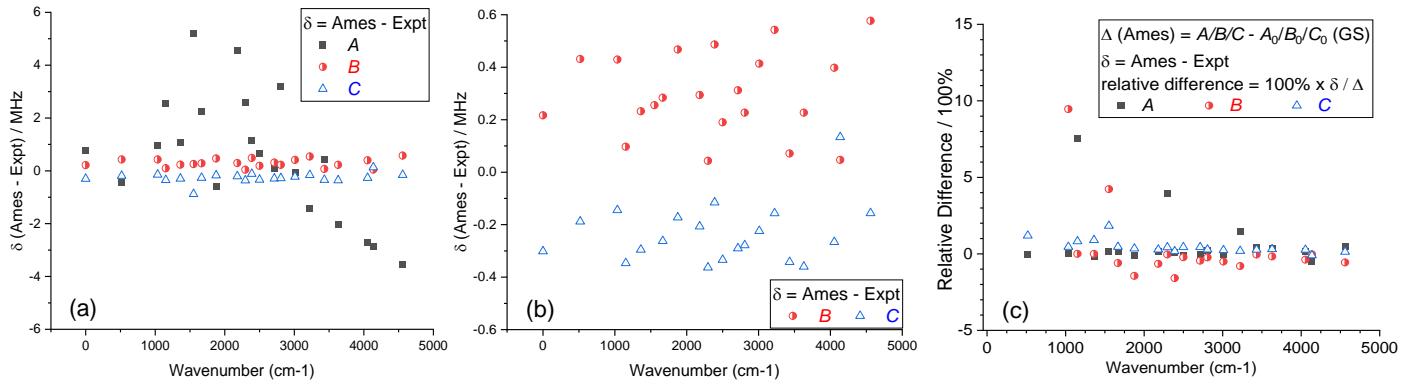


Fig.3 a) The  $\delta_{A/B/C}$  differences on 21  $\text{SO}_2$  626 vibrational states,  $\delta = X_{\text{Ames}} - X_{\text{Expt}}$ ; b) the  $\delta_B$  and  $\delta_C$  differences,  $\delta C(3\nu_2)$  is out of range ; c) the relative differences as defined in text and the plot.

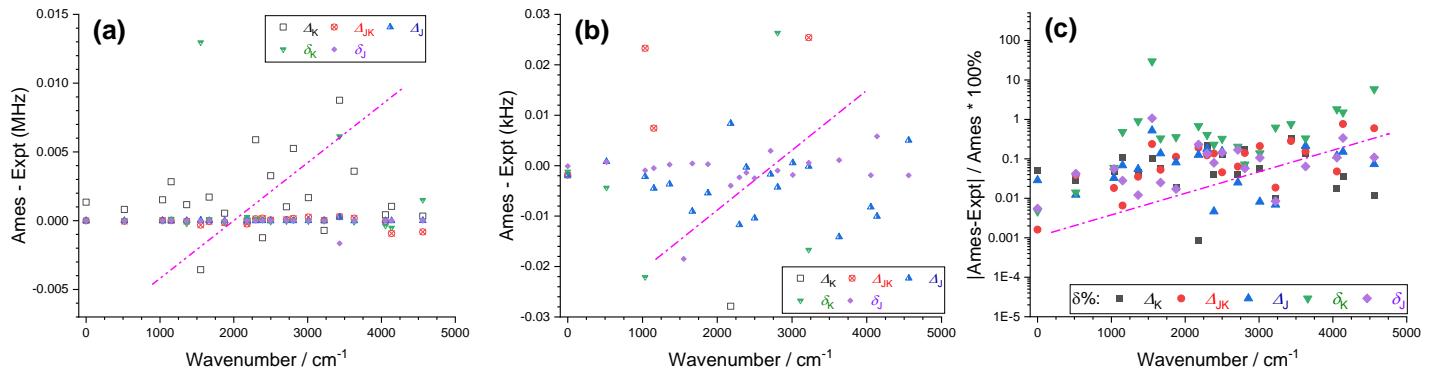


Fig.4 Ames vs Expt differences of five quartic centrifugal distortion constants of  $SO_2$  626 vibrational states: a)-b)  $\delta = X_{Ames} - X_{Expt}$ ; c) relative difference  $\delta\% = 100\% \times \delta/X_{Ames}$ .

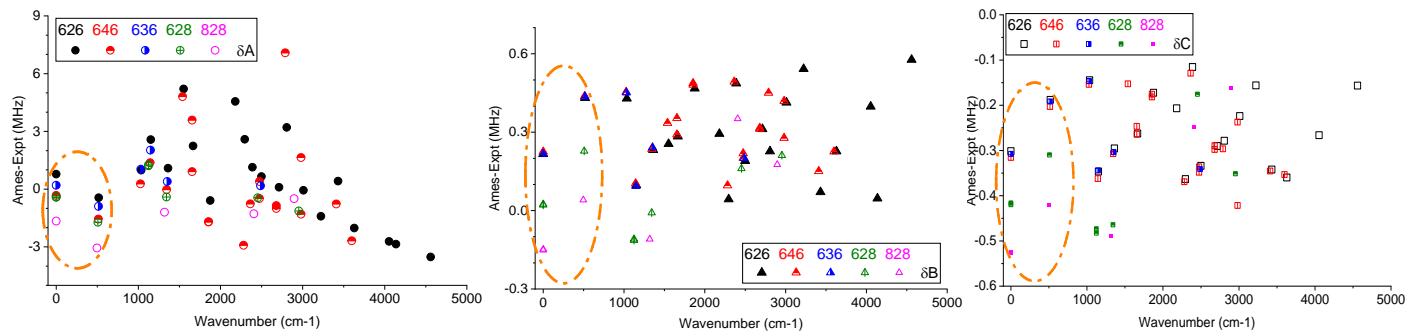


Fig.5 The  $\delta = X_{\text{Ames}} - X_{\text{Expt}}$  differences of A/B/C constants for 5 isotopologues: a) A; b) B; c) C.

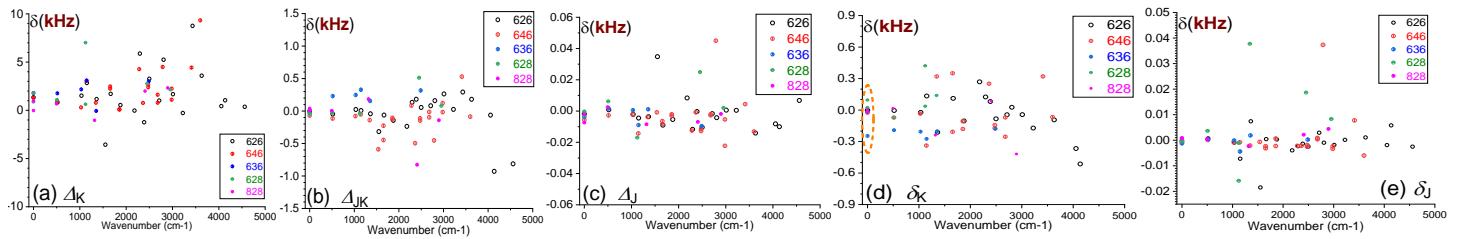


Fig.6  $\delta(X_{\text{Ames}}^{\text{ISO}} - X_{\text{Expt}}^{\text{ISO}})$  for the five quartic EH terms of  $5\text{SO}_2$  isotopologues: a)  $\Delta_K$ ; b)  $\Delta_{JK}$ ; c)  $\Delta_J$ ; d)  $\delta_K$ ; e)  $\delta_J$  in kHz.

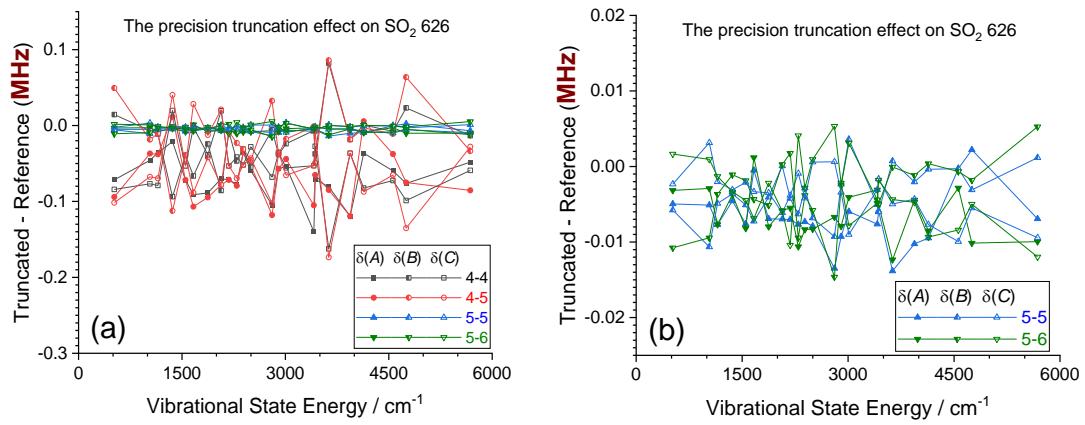


Fig.7 A/B/C differences between the EH(truncated) and EH(1E-7) reference values. a) four truncated sets in a range of 0.3 MHz; b) 1E-5  $\text{cm}^{-1}$  (0.3 MHz) truncation differences in a range of 0.02 MHz.

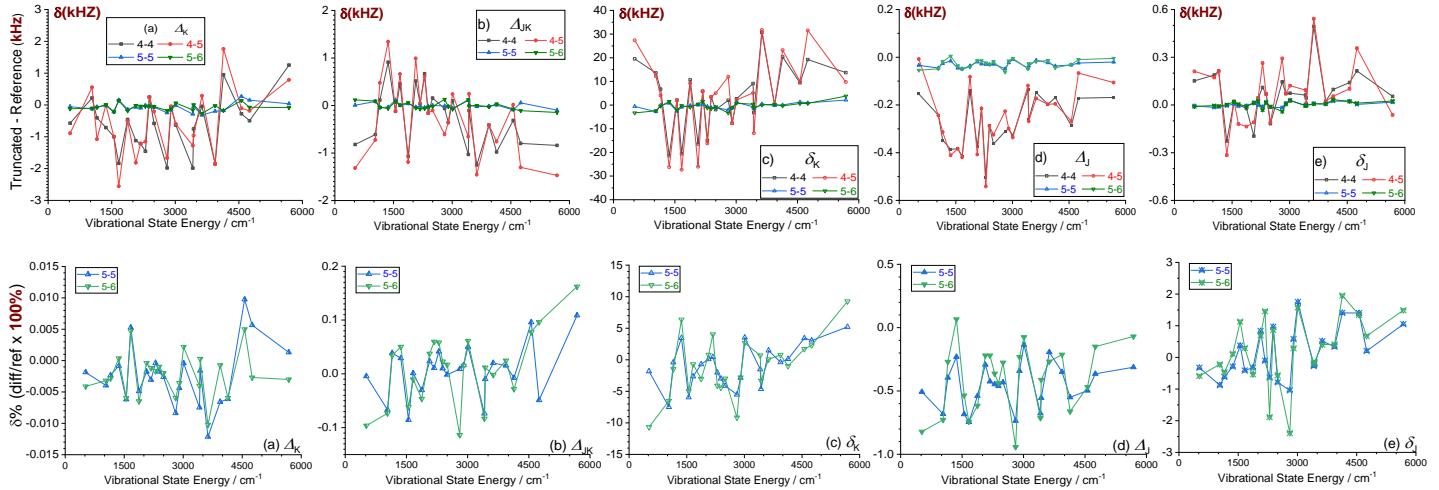


Fig.8  $\delta = X(\text{truncated}) - X(\text{reference 1E-7})$  differences (top row) and relative differences  $\delta\% = \delta / X(\text{reference 1E-7})$  (bottom row) for 5 quartic constants, from 4 truncation tests: a) $\Delta_K$ ; b) $\Delta_{JK}$ ; c) $\delta_K$ ; d) $\Delta_J$ ; e) $\delta_J$